

REQUEST FOR FILING A PATENT APPLICATION UNDER 37 C.F.R. §1.53(b)

Attorney Docket: 147/37315CP

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Box PATENT APPLICATION

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

This is a request for filing a divisional application under 37 C.F.R. 1.53(b), of pending prior application:

Serial No. 08/922,635
Filed on 09/03/97
Of John E. PILETZ et al.
Entitled: DNA MOLECULES ENCODING IMIDAZOLINE
RECEPTIVE POLYPEPTIDES AND POLYPEPTIDES ENCODED THEREBY
Examiner: E. Lazar-Wesley
Group: 1646
Batch No: W07

- X 1. Accompanying this order is a true copy of the prior application as originally filed which includes a specification of 99 pages, 36 claims and 6 sheets of drawings of Figures 1-6.
- X 2. New Formal drawings are being filed herewith consisting of 2 sheet(s), depicting Figures 6A-6B.
- X 3. Declaration and Power of Attorney:
- a. Newly executed (original or copy)
- b. X Copy from prior application. Also attached is a copy of the Substitute Power of Attorney which was filed on July 22, 1999 in prior application Serial No. 08/922,635.
- i. Deletion of Inventors - signed statement attached deleting inventor(s) named in the prior application
- X 4. Incorporation by Reference:

The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 3b, is considered as being part of the disclosure of the accompanying

application and is hereby incorporated by reference therein.

- X 5. Cancel original claims 1-15 and 25-26.
6. Small entity status:
- a. A small entity statement is enclosed
- b. A statement of small entity status (copy attached) was filed on in the prior application and status as a small entity is still proper and desired.
- c. Is no longer claimed
- X 7. The filing fee is calculated below:

CLAIMS AS FILED, INCLUDING ANY CLAIMS
CANCELLED OR ADDED BY PRELIMINARY AMENDMENT

Basic Fee					\$ 760.00
Total Claims	<u>23</u>	- 20 =	<u>3</u>	x 9 =	\$ <u>0</u> 18 = \$ <u>54.00</u>
Ind. Claims	<u>5</u>	- 3 =	<u>2</u>	x 39 =	\$ <u>0</u> 78 = \$ <u>78.00</u>
<u> </u> Multiple Dependent Claims		+ 130 =	\$ <u> </u>	260 =	\$ <u> </u>
		Total		\$ <u> </u>	\$ 892.00

8. Please charge my Deposit Account No. 05-1323 (Docket #) in the amount of \$.
- X 9. A check in the amount of \$ 892.00 to cover the filing fee is enclosed.
- X 10. The Commissioner is authorized to charge any fee which may be required under 37 CFR 1.16 or 37 CFR 1.17 or credit any overpayment to Deposit Account No. 05-1323 (Docket 147/37315D2).
- X 11. Amend the specification by inserting before the first line the sentence: --This application is a division of application Serial No. 09/922,635, filed September 3, 1997, which is a continuation-in-part of application Serial No. 08/650,766, filed May 20, 1996, which is related to provisional application Serial No. 60/12,600, filed March 1, 1996, abandoned.--

- _____ 12. Priority of Appln. No(s). _____, filed in _____
on _____, is hereby claimed under 35 U.S.C.
119.
- _____ 13. A certified copy of each said priority document
was filed in application Serial No. _____ on
_____.
- X 14. The prior application is assigned of record to
The University of Mississippi Medical Center.
- X 15. The Substitute power of attorney in the prior
application is to:

Herbert I. Cantor, Reg. No. 24,392; James F.
McKeown, Reg. No. 25,406; Donald D. Evenson, Reg.
No. 26,160; Joseph D. Evans, Reg. No. 26,269;
Gary R. Edwards, Reg. No. 31,824; Jeffrey D.
Sanok, Reg. No. 32,169; and James M. Verna, Reg.
No. 33,287

- X a. The Substitute power appears in the prior
application.
- _____ b. Since the power does not appear in the
original application papers, a copy of the
power in the prior application is enclosed.
- _____ c. Attached is a duplicate of a Supplemental
Declaration which was filed in the prior
application to overcome informalities.
- X d. Address all future correspondence to:

J. D. Evans
EVENSON, MCKEOWN, EDWARDS
& LENAHAN, P.L.L.C.
1200 G Street, N.W.
Suite 700
Washington, DC 20005

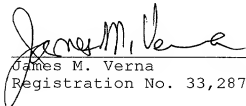
- X 16. Forms PTO-892 and PTO-1449 listing
prior art made of record in the prior
application are attached. A copy of
each of the listed references should be
available in the prior application
file.
- X 17. A Preliminary Amendment is being filed herewith.
- X 18. Return Receipt Postcard.

X 19. Attached is a letter requesting the use of the computer readable form of the Corrected Sequence Listing submitted on May 7, 1999 in pending prior application Serial No. 08/650,766 in this Division Application.

X 20. Other: Corrected Sequence Listing
Substitute Drawing of Figures 6A-6B

Respectfully submitted,

October 8, 1999


James M. Verna
Registration No. 33,287

J.D. Evans
Registration No. 26,269

EVENSON, McKEOWN, EDWARDS
& LENAHA, P.L.L.C.
1200 G Street, N.W., Suite 700
Washington, D.C. 20005
Tel: (202) 628-8800
Fax: (202) 628-8844

JDE:JMV:atc

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: JOHN E. PILETZ ET AL.

Serial No.: Divisional of
08/922,635

Prior Group Art Unit: 1646

Filed: OCTOBER 8, 1999

Prior Examiner: E. Lazar-Wesley

Title: DNA MOLECULES ENCODING IMIDAZOLINE RECEPTIVE
POLYPEPTIDES AND POLYPEPTIDES ENCODED THEREBY

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Please enter the following amendments to the claims prior to the examination of the continued prosecution application.

IN THE CLAIMS:

Please amend Claims 16, 19, 22-24, 27, 32, 35 and 36 as follows:

16. (Amended) An isolated polypeptide including a site which is receptive to imidazoline compounds, said polypeptide containing an amino acid sequence with at least [80] 90% sequence [similarity] identity with the amino acid sequence shown in SEQ ID [No.]NO: 6, wherein the percent identity is determined using the BLASTP program with default parameters.

19. (Amended) An isolated polypeptide including a site which is receptive to imidazoline compounds, said polypeptide containing an amino acid sequence with at least [80] 90% sequence [similarity] identity with the amino acid sequence shown in SEQ ID [No.]NO: 5, wherein the percent identity is determined using the BLASTP program with default parameters.

22. (Amended) A fragment of the amino acid sequence shown in SEQ ID [No.] NO:5 or 6, which fragment is receptive to imidazoline compounds.

23. (Amended) A polypeptide [according to any one] of claim[s] 16 [to 22], which is immunoreactive with at least one of Reis antiserum and Dontenwill antiserum.

24. (Amended) A polypeptide [according to any one] of claim[s] 16 [to 23], which is a human polypeptide.

27. (Amended) A method of screening for a ligand of an imidazoline receptor, which method comprises:

culturing a host cell [as defined in claim 15] transfected with a vector containing an isolated DNA molecule in a culture medium capable of [to] expressing a polypeptide including an amino acid sequence which is receptive to imidazoline compounds;

contacting said polypeptide with a labelled ligand for the imidazoline receptor under conditions effective to bind the labelled ligand thereto;

contacting said polypeptide with a candidate ligand; and
detecting any displacement of the labelled ligand from said polypeptide, wherein displacement signifies that the candidate ligand is a ligand for the imidazoline receptor.

32. (Amended) A method of obtaining a DNA material encoding a polypeptide which is receptive to imidazoline compounds, said method comprising:

providing a labelled DNA probe by labelling a DNA molecule identical or complementary to a DNA molecule comprising a DNA sequence with at least 90% sequence identity with the DNA sequence shown in SEQ ID NO:1, 2, 3, 4 or 5, wherein the percent identity is determined using the BLASTN program with default parameters [as defined in any one of claims 1 to 9] or a [restriction] fragment thereof;

contacting said DNA probe with genetic material suspected of encoding said imidazoline receptive polypeptide;

hybridizing said DNA probe and said genetic material under

stringent hybridization conditions;

identifying any portion of the genetic material which hybridizes to said DNA probe; and

isolating said identified material.

35. (Amended) A method according to claim 32, wherein the labelled DNA probe is provided by labelling a 1110 bp ApaI-EcoRI restriction fragment from a DNA molecule comprising a DNA sequence with at least 90% sequence identity with the DNA sequence shown in SEQ ID NO:1, 2, 3, 4 or 5, wherein the percent identity is determined using the BLASTN program with default parameters [according to claim 12 or 13].

36. (Amended) A method of raising antibodies immunoreactive with a polypeptide which is receptive to an imidazoline compound, which method comprises:

injecting an animal with a polypeptide as defined in [any one of] claim[s] 16 [to 24 and 26]; and

isolating antibodies produced by the animal.

Kindly add new Claims 37-40 as follows:

--37. A polypeptide of claim 19, which is immunoreactive with at least one of Reis antiserum and Dontenwill antiserum.

38. A polypeptide of claim 19, which is a human polypeptide.

39. A method according to claim 32, wherein the labelled DNA probe is provided by labelling a 1.85 kb ApaI-EcoRI restriction fragment from a DNA molecule comprising a DNA sequence with at least 90% sequence identity with the DNA sequence shown in SEQ ID NO:1, 2, 3, 4 or 5 and operably linked with a promoter sequence, wherein the percent identity is determined using the BLASTN program with default parameters.

40. A method of raising antibodies immunoreactive with a polypeptide which is receptive to an imidazoline compound, which

method comprises:

- injecting an animal with a polypeptide as defined in claim 19; and
- isolating antibodies produced by the animal.--

REMARKS

The above amendments have been made to reduce the number of claims prior to calculation of the filing fee and to eliminate the multiple claim dependencies.

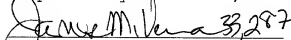
Favorable action on the application is earnestly solicited.

If there are any questions regarding this Preliminary Amendment or this application in general, a telephone call to the undersigned would be appreciated since this should expedite the prosecution of the application for all concerned.

It is respectfully requested that, if necessary to effect a timely response, this paper be considered as a Petition for an Extension of Time sufficient to effect a timely response and shortages in other fees, be charged, or any overpayment in fees be credited, to the Account of Evenson, McKeown, Edwards & Lenahan, P.L.L.C., Deposit Account No. 05-1323 (Docket #147/37315D2).

October 8, 1999

Respectfully submitted,


James M. Verna, Ph.D.
Registration No. 33,287

J. D. Evans
Registration No. 26,269

EVENSON, McKEOWN, EDWARDS
& LENAHA
1200 G Street, N.W., Suite 700
Washington, DC 20005
Telephone No.: (202) 628-8800
Facsimile No.: (202) 628-8844

JDE:JMV:atc

DNA MOLECULES ENCODING IMIDAZOLINE RECEPTIVE POLYPEPTIDES
AND POLYPEPTIDES ENCODED THEREBY

REFERENCE TO RELATED APPLICATION

The present application is a continuation-in-part of application Serial No. 08/650,766 filed May 20, 1996, which is related to provisional application Serial No. 60/12,600, filed March 1, 1996.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention is directed to DNA molecules encoding imidazoline receptive polypeptides, preferably encoding human imidazoline receptive polypeptides, that can be used as an imidazoline receptor (abbreviated IR). In addition, transcript(s) and protein sequences are predicted from the DNA clones. The invention is also directed to a genomic DNA clone designated as JEP-1A. The cDNA clones according to the invention comprise cDNA homologous to portion(s) of this genomic clone; including 5A-1 cDNA, cloned by the inventors that established the open-reading frame for translation of mRNA from the gene, and established the immunoreactive properties of its polypeptide sequence in an expression systems. Also, the invention relates to cDNA clone EST04033, which is another clone identified to contain cDNA sequences from the JEP-1A gene, and of which the 5A-1 is a part, that encodes an active fragment of the IR polypeptide in transfection assays, and the protein sequences thereof. The

invention also relates to methods for producing such genomic and cDNA clones, methods for expressing the IR protein and fragments, and uses thereof.

2. Description of Related Art

5 It is believed that brainstem imidazoline receptors possess binding site(s) for therapeutically relevant imidazoline compounds, such as clonidine and idazoxan. These drugs represent the first generation of ligands discovered for the binding site(s) of imidazoline receptors. However, clonidine and
0 idazoxan were developed based on their high affinity for α_1 -adrenergic receptors. Second generation ligands, such as moxonidine, possess somewhat improved selectivity for IR over α_2 -adrenergic receptors, but more selective compounds for IR are needed.

5 An imidazoline receptor clone is of particular interest because of its potential utility in identifying novel pharmaceutical agents having greater potency and/or more selectivity than currently available ligands have for imidazoline receptors. Recent technological advances permit pharmaceutical
0 companies to use combinatorial chemistry techniques to rapidly screen a cloned receptor for ligands (drugs) binding thereto. Thus, a cloned imidazoline receptor would be of significant value to a drug discovery program.

Until now, the molecular nature of imidazoline receptors
5 remains unknown. For instance, no amino acid sequence data for a novel IR, e.g., by N-terminal sequencing, has been reported.

Three different techniques have been described in the literature by three different laboratories to visualize imidazoline-selective binding proteins (imidazoline receptor candidates) using gel electrophoresis. Some important consistencies have emerged from these results despite the diversity of the techniques employed. On the other hand, multiple protein bands have been identified, which suggests heterogeneity amongst imidazoline receptors. These reports are discussed below.

Some of the abbreviations used hereinbelow, have the following meanings:

α_2 AR	Alpha-2 adrenoceptor
BAC	Bovine adrenal chromaffin
ECL	Enhanced chemiluminescence (protein detection procedure)
EST	Expressed Sequence Tag (a one-pass cDNA documentation without identification)
I-site	Any imidazoline-receptive binding site (e.g., encoded on IR)
IR ₁	Imidazoline receptor subtype ₁
IR-Ab	Imidazoline receptor antibody
I ₂ Site	Imidazoline binding subtype ₂
kDa	Kilodaltons (molecular size)
MAO	monoamine oxidase
MW	molecular weight
NRL	European abbreviation for RVLM (see below)
PC-12	Phaeochromocytoma-12 cells
¹²⁵ I PIC	[¹²⁵ I]p-iodoclonidine
PKC	Protein Kinase C
RVLM	Rostral Ventrolateral Medulla in brainstem
SDS	sodium dodecyl sulfate gel electrophoresis

Reis et al. [Wang et al., Mol. Pharm., 42: 792-801 (1992); Wang et al., Mol. Pharm., 43: 509-515 (1993)] were the first to characterize an imidazoline-selective binding protein and to demonstrate it as having MW = 70 kDa. This was accomplished using bovine cells (BAC), which lack an α_2 AR [Powis & Baker, Mol.

Pharm., 29:134-141 (1986)]. The 70 kDa imidazoline-selective protein in those studies had high affinities for both idazoxan and p-aminoclonidine affinity chromatography columns and was eluted by another imidazoline compound (phentolamine). Unfortunately, those investigators failed to isolate sufficient 70 kDa protein to determine its other biochemical properties. To date, no one has reported the complete purification of an imidazoline receptor protein. Likewise, no amino acid sequences have been reported for IR.

Their 70 kDa protein was used by Reis and co-workers to raise "I-site binding antiserum", designated herein as Reis antiserum. The term "I-site" refers to the imidazoline binding site, presumably defined within the imidazoline receptor protein. Reis antiserum was prepared by injecting the purified protein into rabbits [Wang et al, 1992]. The first immunization was done subcutaneously with the protein antigen (10 μ g) emulsified in an equal volume of complete Freund's adjuvant, and the next three booster shots were given at 15-day intervals with incomplete Freund's adjuvant. The polyclonal antiserum has been mostly characterized by immunoblotting, but radioimmunoassays (RIA) and/or conjugated assay procedures, i.e., ELISA assays, are also conceivable [see "Radioimmunoassay of Gut Regulatory Peptides: Methods in Laboratory Medicine," Vol. 2, chapters 1 and 2, Praeger Scientific Press, 1982].

The present inventors and others [Escriba et al., Neurosci. Lett. 178: 81-84 (1994)] have characterized the Reis antiserum in

several respects. For instance, the present inventors have discovered that human platelet immunoreactivity with Reis antiserum is mainly confined to a single protein band of MW \approx 33 kDa, although a trace band at \approx 85 kDa was also observed. The \approx 33 and \approx 85 kDa bands were enriched in plasma membrane fractions as expected for an imidazoline receptor. Furthermore, the intensity of the \approx 33 kDa band was found to be positively correlated with non-adrenergic 125 PIC Bmax values at platelet IR₁ sites in samples from the same subjects, with an almost one-to-one slope factor. In addition, the nonadrenergic 125 PIC binding sites on platelets were discovered by the present inventors to have the same rank order of affinities as IR₁ binding sites in brainstem [Piletz and Sletten, J. Pharm. & Exper. Therap., 267: 1493-1502 (1993)]. The platelet \approx 33 kDa band may also be a product of a larger protein, since in human megakaryoblastoma cells, which are capable of forming platelets in tissue cultures, an \approx 85 kDa immunoreactive band was found to predominate.

Immunoreactivity with Reis antiserum does not appear to be directed against human α_2 AR and/or MAO A/B. This is a significant point because α_2 AR and MAO A/B have previously been cloned and also bind to imidazolines. The present inventors have obtained selective antibodies and recombinant preparations for α_2 AR and MAO A/B, and these proteins do not correspond to the \approx 33, 70, or 85 kDa putative IR₁ bands. Thus, there is substantial evidence that, at least in human platelets, the Reis antiserum is

IR₁ selective.

Another antiserum was raised by Drs. Dontenwill and Bousquet in France [Grenay et al., Europ. J. Pharmacol., 265: R1-R2 (1994); Grenay et al., Neurochem. Int., 25: 183-191 (1994); Bennai et al., Annals NY Acad. Sci., 763:140-148 (1995)] against polyclonal antibodies for idazoxan (designated Dontenwill antiserum). This anti-idiotypic antiserum inhibits ³H-clonidine but not ³H-rauwolscine (α_2 -selective) binding sites in the brainstem, suggesting it also interacts with IR₁ [Bennai et al., 1995]. As shown in Fig. 1, human RVLM (same as NRL) membrane fractions displayed bands of \approx 41 and 44 kDa, as detected by the present inventors using this anti-idiotypic antiserum.

The present inventors have found that the bands of MW \approx 41 and 44 kDa detected by Dontenwill antiserum may be derived from an \approx 85 kDa precursor protein, similar to that occurring in platelet precursor cells. An 85 kDa immunoreactive protein is obtained in fresh rat brain membranes only when a cocktail of 11 protease inhibitors is used. Also, as shown in Fig. 1, it is found that Reis antiserum detects the \approx 41 and 44 kDa bands in human brain when fewer protease inhibitors are used. Additionally, the Dontenwill antiserum weakly detects a platelet \approx 33 kDa band. Thus, the present inventors have hypothesized that the \approx 41 and 44 kDa immunoreactive proteins may be alternative breakdown products of an \approx 85 kDa protein, as opposed to the platelet \approx 33 kDa breakdown product.

In summary, the main conclusion from the above results is

that, despite vastly different origins, the Reis and Dontenwill antisera both detect identical bands in human platelets, RVLM, and hippocampus.

Using yet another technique, a photoaffinity imidazoline ligand, ¹²⁵AZIPI, has also been developed to preferentially label I₂-imidazoline binding sites [Lanier et al., J.Biol.Chem., 268: 16047-16051 (1993)]. The ¹²⁵AZIPI photoaffinity ligand was used to visualize \approx 55 kDa and \approx 61 kDa binding proteins from rat liver and brain. It is believed that the \approx 61 kDa protein is probably MAO, in agreement with other findings [Tesson et al., J.Biol.Chem., 270: 9856-9861 (1995)] showing that MAO proteins bind certain imidazoline compounds. The different molecular weights between these bands and those detected immunologically by the present inventors is one of many pieces of evidence that distinguishes IR₁ from I₂ sites.

To the inventors' knowledge and as described herein, we are first to clone the gene, cDNAs and fragments thereof encoding a protein with the immunological and ligand binding properties expected of an IR. On this basis, we are first to identify the nucleotide sequences of DNA molecules encoding an imidazoline receptor and active fragments thereof, and the first to determine the amino acid sequence of an imidazoline receptor and active fragments thereof. The polypeptides described herein are clearly distinct from α_2 AR or MAO A/B proteins.

SUMMARY OF THE INVENTION

The present invention involves various cDNA clones (ie., 5A-1 and EST04033) and a genomic clone (JEP-1A) which are directed to an isolated polypeptide(s) that is receptive to (bind to) imidazoline compound(s), and can be used to identify other compounds of interest. Currently available imidazoline compounds in this context are p-iodoclonidine and moxonidine. Initially, the inventors detected a polypeptide expressed by their cDNA clone (5A-1 isolated from a human hippocampus cDNA library) that immunoreacted with Reis antiserum and/or Dontenwill antiserum. The DNA sequence of the 5A-1 clone is encapsulated within a portion of the other clones (EST04033 and JEP-1A genomic clone).

In one aspect of the invention, a polypeptide includes a 651 amino acid sequence as shown in SEQ ID No. 5. This polypeptide is predicted from non-plasmid cDNA in EST04033; a clone which the inventors showed possesses sequences inclusive of 5A-1. Furthermore, transfection of EST04033 into COS cells yielded imidazoline receptivity by radioligand binding assays (described in detail later). Other imidazoline receptive proteins homologous to this polypeptide are also contemplated. Such polypeptide(s) generally have a molecular weight of about 50 to 80 kDa. More particularly, one can have a molecular weight of about 70 kDa.

In another aspect of this invention, a polypeptide includes a 390 amino acid sequence as shown in SEQ ID No. 6. This represents the polypeptide predicted from the non-plasmid DNA of

the original 5A-1 clone. Such a polypeptide generally has a molecular weight of about 35 to 50 kDa. More particularly, it can have a molecular weight of about 43 kDa.

DNA molecules encoding aforementioned imidazoline-receptive polypeptide(s) are also contemplated. Such a DNA molecule, e.g., a cDNA derived from mRNA, can contain a nucleotide sequence encoding the 651 amino acid sequence shown in SEQ ID No. 5. Thus, a DNA molecule containing the 1954 base pairs (b.p.) (1954 b.p. encodes 651 amino acids) nucleotide sequence shown in SEQ ID No. 2 is contemplated. This represents the coding sequence for the polypeptide predicted by EST04033 transfections. In another embodiment, a DNA molecule includes the longer nucleotide sequence shown in SEQ ID No. 3. This represents the cDNA predicted to have been translated + not predicted to have been translated in transfections experiments of EST04033.

In another embodiment of the invention, a DNA molecule contains a nucleic acid sequence encoding the amino acid sequence shown in SEQ ID No. 6. In another aspect, it can include the 1171 b.p. nucleic acid sequence shown in SEQ ID No. 4. The 1171 b.p. nucleic acid sequence shown in SEQ ID No. 4 is the 5A-1 non-plasmid DNA.

The nucleic acid sequence of the genomic clone encoding the imidazoline receptor is further shown in SEQ ID No. 21. The nucleic acid and amino acid sequence of the predicted transcript (ie., cDNA) can be predicted from the description hereinbelow. The polypeptide encoded by the genomic DNA is shown in SEQ ID No.

22.

Sequence similarity with the sequences indicated in SEQ ID protocols of the attached Sequence Listing is defined in connection with the present invention as a very close structural relationship of the relevant sequences with the sequences indicated in the respective SEQ ID protocols. To determine the sequence similarity, in each case the structurally mutually corresponding sections of the sequence of the SEQ ID protocol and of the sequence to be compared therewith are superimposed in such a way that the structural correspondence between the sequences is a maximum, account being taken of differences caused by deletion or insertion of individual sequence members (DNA-codon or amino acid respectively), and being compensated by appropriate shifts in sections of the sequences. The sequence similarity in % results from the number of sequence members which now correspond to one another in the sequences ("homologous positions") relative to the total number of members contained in the sequences of the SEQ ID protocols. Differences in the sequences may be caused by variation, insertion or deletion of sequence members.

Additionally in DNA sequences, different DNA-codons encoding for the same amino acid are considered identical in the context of the present invention. For amino acid sequences, conservative amino acid substitutions encoded by their corresponding DNA-codons, as well as naturally occurring homologs of the sequences, are considered within the context of sequence similarity.

DNA molecules of substantial homology ($\geq 75\%$) are an

implicit aspect of this sort of invention. As will be discussed later, the inventors have already identified two possible splice variants in the amino acid coding sequence. In addition, artificially mutated receptor cDNA molecules can be routinely constructed by methods such as site-directed polymerase chain reaction-mediated mutagenesis [Nelson and Long, Anal. Biochem. 180: 147-151 (1989)]. It is commonly appreciated that highly homologous mutants frequently mimic their natural receptor. A study by Kjelsberg et al. [J. Biol. Chem. 267: 1430-1433 (1992)] showed that all 20 amino acid substitutions produce an active receptor at a single site in the α_{1B} -adrenergic receptor. RNA molecules of $\geq 75\%$ complementarity to an instant DNA molecule, e.g., an mRNA molecule (sense) or a complementary cRNA molecule (antisense), are a further aspect of the invention.

A further aspect of the invention is for a recombinant vector, as well as a host cell transfected with the recombinant vector, wherein the recombinant vector contains at least one of the nucleotide sequences shown in SEQ ID Nos. 1-4, or sequences predicted by the genomic clone, or nucleotide sequences $\geq 75\%$ homologous thereto.

A method of producing an imidazoline receptor protein is another aspect of the invention. Such a method entails transfecting a host cell with an aforementioned vector, and culturing the transfected host cell in a culture medium to generate the imidazoline receptor.

A method for producing homologous imidazoline receptor

proteins, and the proteins produced thereby, are also considered an aspect of this invention.

A significant further aspect of the invention is a method of screening for a ligand that binds to an imidazoline receptor. Such a method can comprise culturing an above-mentioned transfected cell in a culture medium to express imidazoline receptor proteins, followed by contacting the proteins with a labelled ligand for the imidazoline receptor under conditions effective to bind the labelled ligand thereto. The imidazoline receptor proteins can then be contacted with a candidate ligand, and any displacement of the labelled ligand from the proteins can be detected. Displacement of labelled ligand signifies that the candidate ligand is a ligand for the imidazoline receptor. These steps could be performed on intact host cells, or on proteins isolated from the cell membranes of the host cells.

The invention will now be described in more detail with reference to specific examples.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 depicts a comparison of Reis antiserum (lane 1, 1:2000 dilution) and Dontenwill antiserum (lane 2, 1:5000 dilution) immunoreactivities for human NRL (same as RVLM) and hippocampus, as discussed in Example 1.

Fig. 2 depicts a comparison of Reis antiserum (1:15,000 dilution) and Dontenwill antiserum (1:20,000 dilution) immunoreactivities for plaques isolated from the human

hippocampal cDNA library used in cloning as discussed in Example 2. The plaques contain the initial clone, designated herein as 5A-1, in a third stage of purification.

Fig. 3 depicts the restriction map of the EST04033 cDNA clone.

Fig. 4 depicts a competitive binding assay between ^{125}I -labelled p-iodoclonidine (PIC) and various ligands for the imidazoline receptor on membranes expressed in COS cells transfected with the EST04033 cDNA clone, as discussed in Example 4.

Fig. 5 depicts the prediction of introns and exons of the genomic clone (as analyzed by the GENESCAN program and verified by the available CDNAS).

Fig. 6 depicts the distribution of mRNA homologous to our CDNA in human adult tissues (bar graph) and the two species of MRNA (6 and 9.5 kb).

DETAILED DESCRIPTION OF THE INVENTION

The present invention is concerned with multiple aspects of an imidazoline receptor protein, and DNA molecules encoding the same, and fragments thereof, which have now been discovered.

First, a polypeptide having imidazoline binding activity has been identified, which contains the putative active site for binding, as discussed hereinafter. Although polypeptide(s) described herein has a binding affinity for an imidazoline compound, it may also have an enzymatic activity, such as do

catalytic antibodies and ribozymes. In fact, one such domain within our protein predicts a cytochrome p450 activity (described later).

Exemplary "binding" polypeptides are those containing either of the amino acid sequences shown in SEQ ID Nos. 5 or 6 (with the amino acid sequence predicted by EST04033 given in SEQ ID No. 5). Functionally equivalent polypeptides are also contemplated, such as those having a high degree of homology with such aforementioned polypeptides, particularly when they contain the Glu-Asp-rich region described hereinafter which we believe may define an active imidazoline binding site.

A polypeptide of the invention can be formed by direct chemical synthesis on a solid support using the carbodiimide method [R. Merrifield, JACS, 85: 2143 (1963)]. Alternatively, and preferably, an instant polypeptide can be produced by a recombinant DNA technique as described herein and elsewhere [e.g., U.S. Patent No. 4,740,470 (issued to Cohen and Boyer), the disclosure of which is incorporated herein by reference], followed by culturing transformants in a nutrient broth.

Second, a DNA molecule of the present invention encodes aforementioned polypeptide. Thus, any of the degenerate set of codons encoding an instant polypeptide is contemplated. A particularly preferred coding sequence is the 1954 b.p. sequence set forth in SEQ ID No. 2, which has now been discovered to be a nucleotide sequence that encodes a polypeptide capable of binding imidazoline compound(s). In another embodiment, a DNA molecule

includes the 3318 b.p. nucleotide sequence shown in SEQ ID No. 3. This latter sequence is the entire EST04033 insert. It includes the nucleotide sequence of SEQ ID No. 2 which was predicted to have been translated into protein in the transfection experiments.

In another embodiment of the invention, a DNA molecule contains a nucleic acid sequence encoding the amino acid sequence (390 residues) shown in SEQ ID No. 6. This amino acid sequence corresponds to that derived from direct sequencing of the 5A-1 clone and represents a fragment of the native protein. The 5A-1 DNA molecule is defined by the 1171 b.p. nucleic acid sequence shown in SEQ ID No. 4.

A DNA molecule of the present invention can be synthesized according to the phosphotriester method [Matteucci et al., JACS, 103: 3185 (1988)]. This method is particularly suitable when it is desired to effect site-directed mutagenesis of an instant DNA sequence, whereby a desired nucleotide substitution can be readily made. Another method for making an instant DNA molecule is by simply growing cells transformed with plasmids containing the DNA sequence, lysing the cells, and isolating the plasmid DNA molecules. Preferably, an isolated DNA molecule of the invention is made by employing the polymerase chain reaction (PCR) [e.g., U.S. Patent No. 4,683,202 (issued to Mullis)] using synthetic primers that anneal to the desired DNA sequence, whereby DNA molecules containing the desired nucleotide sequence are amplified. Also, a combination of the above methods can be

employed, such as one in which synthetic DNA is ligated to CDNA to produce a quasi-synthetic gene [e.g., U.S. Patent No. 4,601,980 (issued to Goeddel et al.)].

A further aspect of the invention is for a vector, e.g., a plasmid, that contains at least one of the nucleotide sequences shown in SEQ ID Nos. 1-4 or those predicted by the genomic clone in SEQ ID No. 21. Whenever the reading frame of the vector is appropriately selected, the vector encodes an IR polypeptide of the invention. Hence, as well as full-length protein, fragments of the native IR protein are contemplated; as well as fusion proteins that incorporate an amino acid sequence as described herein. Also, a vector containing a nucleotide sequence having a high degree of homology with any of SEQ ID Nos. 1-4 or 21 is contemplated within the invention, particularly when it encodes a protein having imidazoline binding activity.

A recombinant vector of the invention can be formed by ligating an afore-mentioned DNA molecule to a preselected expression plasmid, e.g., with T4 DNA ligase. Preferably, the plasmid and DNA molecule are provided with cohesive (overlapping) termini, with the plasmid and DNA molecule operatively linked (i.e., in the correct reading frame).

Another aspect of the invention is a host cell transfected with a vector of the invention. Relatedly, a protein expressed by a host cell transfected with such a vector is contemplated, which protein may be bound to the cell membrane. Such a protein can be identical with an aforementioned polypeptide, or it can be

a fragment thereof, such as when the polypeptide has been partially digested by a protease in the cell. Also, the expressed protein can differ from an aforementioned polypeptide, as whenever it has been subjected to one or more post-translational modifications. For the protein to be useful within the context of the present invention, it should exhibit imidazoline binding capacity.

A method of producing an imidazoline receptor protein is another aspect of the invention, which entails transfecting a host cell with an aforementioned vector, and culturing the transfected host cell in a culture medium to generate the imidazoline receptor. The receptor molecule can undergo any post-translational modification(s), including proteolytic decomposition, whereby its structure is altered from the basic amino acid residue sequence encoded by the vector. A suitable transfection method is electroporation, and the like.

With respect to transfecting a host cell with a vector of the invention, it is contemplated that a vector encoding an instant polypeptide can be transfected directly in animals. For instance, embryonic stem cells can be transfected, and the cells can be manipulated in embryos to produce transgenic animals. Methods for performing such an operation have been previously described [Bond et al., Nature, 374:272-276 (1995)]. These methods for expressing an instant CDNA molecule in either tissue culture cells or in animals can be especially useful for drug discovery.

Possibly the most significant aspect of the present invention is in its potential for affording a method of screening for a ligand (drug) that binds to an imidazoline receptor. Such a method comprises culturing an above-mentioned host cell in a culture medium to express an instant imidazoline receptive polypeptide, then contacting the polypeptides with a labelled ligand, e.g., radiolabelled p-iodoclonidine, for the imidazoline receptor under conditions effective to bind the labelled ligand thereto. The polypeptides are further contacted with a candidate ligand, and any displacement of the labelled ligand from the polypeptides is detected. Displacement signifies that the candidate ligand actually binds to the imidazoline receptor. These steps could be performed on intact host cells, or on proteins isolated from the cell membranes of the host cells.

Typically, a suitable drug screening protocol involves preparing cells (or possibly tissues from transgenic animals) that express an instant imidazoline receptive polypeptide. In this process, categories of chemical structure are systematically screened for binding affinity or activation of the receptor molecule encoded by the transfected CDNA. This process is currently referred to as combinatorial chemistry. With respect to the imidazoline receptor, a number of commercially available radioligands, e.g., ^{125}PIC , can be used for competitive drug binding affinity screening.

An alternative approach is to screen for drugs that elicit or block a second messenger effect known to be coupled to

activation of the imidazoline receptor, e.g., moxonidine-stimulated arachidonic acid release. Even with a weak binding affinity or activation by one category of chemicals, systematic variations of that chemical structure can be studied and a preferred compound (drug) can be deduced as being a good pharmaceutical candidate. Identification of this compound would lead to animal testing and upwards to human trials. However, the initial rationale for drug discovery becomes vastly improved with an instant cloned imidazoline receptor.

Along these lines, a drug screening method is contemplated in which a host cell of the invention is cultured in a culture medium to express an instant imidazoline receptive polypeptide. Intact cells are then exposed to an identified agent (ie., agonist, inverse agonist, or antagonist) under conditions effective to elicit a second messenger or other detectable responses upon interacting with the receptor molecule. The imidazoline receptive polypeptides are then contacted with one or more candidate chemical compounds (drugs), and any modification in a second messenger response is detected. Compounds that mimic an identified agonist would be agonist candidates, and those producing the opposite response would be inverse agonist candidates. Those compounds that block the effects of a known agonist would be antagonist candidates for an in vivo imidazoline receptor. For meaningful results, the contacting step with a candidate compound is preferably conducted at a plurality of candidate compound concentrations.

A method of probing for another gene encoding an imidazoline receptor or homologous protein is further contemplated. Such a method comprises providing a radiolabelled DNA molecule identical or complementary to one of the above-described cDNA molecules (probe). The probe is then placed in contact with genetic material suspected of containing a gene encoding an imidazoline receptor or encoding a homologous protein, under stringent hybridization conditions (e.g., a high stringency wash condition is $0.1 \times \text{SSC}$, 0.5% SDS at 65°C), and identifying any portion of the genetic material that hybridizes to the DNA molecule.

Still further, a method of selectively producing antibodies, (e.g., monoclonal antibodies, immunoreactive with an instant imidazoline-receptive protein) comprises injecting a mammal with an aforementioned polypeptide, and isolating the antibodies produced by the mammal. This aspect is discussed in more detail in an example presented hereinafter.

The present inventors began their search for a human imidazoline receptor cDNA by screening a $\lambda\text{gt}11$ phage human hippocampus cDNA expression library. Their research had indicated that both of the known antisera (Reis and Dontenwill) that are directed against human imidazoline receptors were immunoreactive with identical bands on SDS gels of membranes prepared from the human hippocampus (an in other tissues). By contrast, other brain regions either were commercially unavailable as cDNA expression libraries or yielded inconsistencies between the two antisera. Therefore, it was felt

that a human hippocampal cDNA library held the best opportunity for obtaining a CDNA for an imidazoline receptor.

Immunoexpression screening was chosen over other cloning strategies because of its sensitivity when coupled with the ECL detection system used by the present inventors, as discussed hereinbelow.

A number of unique discoveries led to identifying the first 5A-1 clone as an imidazoline receptor CDNA. These included discoveries that led to the choice of a hippocampal CDNA library and adapting ECL to the antisera. Once the initial clone (5A-1) was identified and sequenced, a more complete clone (EST04033) was purchased without restriction from ATCC Inc. (Catalogue # 82815; American Type Culture Collection, Rockville, MD). EST 04033 was the only EST clone available at the time of the discovery of 5A-1, that contained a segment of complete homology (the origination of EST04033 is discussed later on). The binding affinities of the expressed protein after transfection in COS cells were determined by radioligand binding procedures developed in the inventor's laboratory [Piletz and Sletten, 1993, *ibid*].

To identify an instant CDNA clone encoding an imidazoline receptor it was preferable to employ both of the known antibodies to imidazoline receptors. These antibodies were obtained by contacting Dr. D. Reis (Cornell University Medical Center, New York City), and Drs. M. Dontenwill and P. Bousquet (Laboratoire de Pharmacologie Cardiovasculaire et Renale, CNRS, Strasbourg, France). These antisera were obtained free of charge and without

confidentiality or restrictions on their use. The former antiserum (Reis antiserum) was derived from a published imidazoline receptor protein [Wang et al., (1992, 1993), the disclosures of which are incorporated herein by reference]. The method for raising the latter antiserum (Dontenwill antiserum) has also been published [Bennai et al., (1995), the disclosure of which is also incorporated herein by reference]. The latter antiserum was developed using an anti-idiotypic approach that identified the pharmacologically correct (clonidine and idazoxan selective) binding site structure.

Example 1. Selectivity of the Antisera.

The obtained Reis antiserum had been prepared against a purified imidazoline binding protein isolated from BAC cells, which protein runs in denaturing-SDS gels at 70 Kda [Wang et al., 1992, 1993]. The Dontenwill antiserum is anti-idiotypic, and thus is believed to detect the molecular configuration of an imidazoline binding site domain in any species. Prior to being used for screening plaques, both antisera were cleaned by stripping out possible antibacterial antibodies.

Both antisera have been tested to ensure that they are in fact selective for a human imidazoline receptor. In particular, we found that both of these antisera detected identical bands in human platelets and hippocampus, and in brainstem RVLM (NRL) by Western blotting (see Fig. 1). In these studies, in order to increase sensitivity over previously published detection methods,

an ECL (Enhanced Chemiluminescence) system was employed. The linearity of response of the ECL system was demonstrated with a standard curve. ECL detection was demonstrated to be very quantifiable and about ten times more sensitive than other screening methods previously used with these antisera. Western blots with antiserum dilutions of 1:3000 revealed immunoreactivity with as little as 1 ng of protein from a human hippocampal homogenate by dot blot analysis.

For the studies depicted in Fig. 1, human hippocampal homogenate (30 μ g) and NRL membrane proteins (10 μ g) were electrophoresed through a 12.5% SDS-polyacrylamide gel, electrotransferred to nitrocellulose and sequentially incubated with (1) the Reis antibody (1:2000 dilution) and (2) the Dontenwill antibody (1:5000 dilution). Immunoreactive bands were visualized with an Enhanced Chemiluminescence (ECL) detection kit (Amersham) using anti-rabbit Ig-HRP conjugated antibody at a dilution of 1:3000 and the ECL detection reagents. Following detection with the antibody, blots were stripped and reprocessed omitting the primary antibody to check for complete removal of this antibody. In panels A and B, lane 1 shows the immunoreactive bands observed with the Reis antibody and lane 2 shows the bands detected with the Dontenwill antibody. Protein molecular weight standards are indicated to the left of each panel (in Kda).

Despite the diverse origins of the Reis and Dontenwill antisera, both of these antisera detected a similar 85 Kda

protein in human brain and other tissues. But, a 33 Kda band was found in human platelets. Although the 33 Kda band is of smaller size than that reported for other tissues [Wang et al., 1993; Escriba et al., 1994; Grenay et al., 1994], the fact that both antisera detected it, suggests that both the 85 Kda and 33 Kda bands may be imidazoline binding polypeptides. The 85 and 33 Kda bands were enriched in plasma membrane fractions, as is known to be the case for IR₁ binding, but not I₂ binding [Piletz and Sletten, 1993].

A significant positive correlation was established for the 85 Kda band detected by the Dontenwill antiserum with IR₁ Bmax values across nine rat tissues ($r^2 = 0.8736$). A similar positive correlation was established amongst platelet samples from 15 healthy platelet donors between radioligand IR₁ Bmax values (but not I₂ or α_2 AR Bmax values), and the 33 Kda band (presumed IR₁ immunoreactivity) on Western blots. This correlation exhibited a slope factor close to unity (results not shown). These correlations strongly suggested that an IR₁ binding protein might be revealed in an imidazoline receptor-antibody Western blotting assay. Furthermore, the Reis antiserum failed to detect authentic α_2 AR, MAO A, or MAO B bands on gels, i.e., it was not immunoreactive with MAO at MW = 61 Kda, or α_2 AR at MW = 64 Kda. Additionally, no immunoreactive bands were observed using preimmune antiserum. Thus, after extensively characterizing these antisera with human and rat materials, it was concluded that these antisera are indeed selective for human

imidazoline receptor protein.

Example 2. Cloning of cDNA For An Imidazoline Receptor

A commercially available human hippocampal cDNA λ gt11 expression library (Clontech Inc., Palo Alto, CA) was screened for immunoreactivity sequentially using both the anti-idiotypic Dontenwill antiserum and the Reis antiserum. Standard techniques to induce protein and transference to a nitrocellulose overlay were employed. [See, for instance, Sambrook et al., 1989, "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press]. After washing and blocking with 5% milk, the Dontenwill antiserum was added to the overlay at 1:20,000 dilution in Tris-buffered saline, 0.05% Tween20, and 5% milk. The Reis antiserum was employed similarly, but at 1:15,000 dilution. These high dilutions of primary antiserum were chosen to avoid false positives. The secondary antibody was added, and positive plaques were identified using ECL. Representative results are shown in Fig. 2.

Positive plaques were pulled and rescreened until tertiary screenings yielded only positive plaques. Four separate positive plaques were identified from more than 300,000 primary plaques in our library. Recombinant λ gt11 DNA purified from each of the four plaques was subsequently subcloned into *E. coli* pBluescript vector (Stratagene, La Jolla, CA). Sequencing of these four cDNA inserts in pBluescript demonstrated that they were identical, suggesting that only one cDNA had actually been identified four

times. Thus, the screening had been verified as being highly reproducible and the frequency of occurrence was as expected for a single copy gene, i.e., one in 75,000 transcripts. As shown in Fig. 2, the protein produced by the first positive clone,

designated 5A-1, tested positive with both the Reis antiserum and the Dontenwill antiserum. Clone 5A-1 has been deposited under the Budapest Treaty with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, MD, USA, 20852, on August 28, 1997 and has been assigned deposit accession no. ATCC 209217.

Tertiary-screened plaques of 5A-1 were all immuno-positive with either of the two known anti-imidazoline receptor antisera, but not with either preimmune antisera. These results suggested that clone 5A-1 encoded a fusion peptide similar to or identical with one of the predominant bands detected in human Western blots by both the Dontenwill and Reis antisera.

Sequencing of the first four clones was performed by contracting with ACGT Company (Chicago, IL) after subcloning them into pBluescript vector SK (Stratagene). Both manual and automatic sequencing strategies were employed which are outlined as follows:

Manual Sequencing

1. DNA sequencing was performed using T7 DNA polymerase and the dideoxy nucleotide termination reaction.

2. The primer walking method [Sambrook et al., ibid.] was used in both directions.

3. (³⁵S)dATP was used for labelling.

4. The reactions were analyzed on 6% polyacrylamide wedge or non-wedge gels containing 8 M urea, with samples being loaded in the order of A C G T.

5. DNA sequences were analyzed by MacVector Version 5.0. and by various Internet-available programs, i.e., the BLAST program.

Automatic Sequencing

1. DNA sequencing was performed by the fluorescent dye terminator labelling method using AmpliTaq DNA polymerase (Applied Biosystems Inc., Prizm DNA Sequencing Kit, Perkin-Elmer Corp., Foster City, CA).

2. The primer walking method was used. The primers actually used were a subset of those shown in SEQ ID Nos. 7-20.

3. Sequencing reactions were analyzed on an Applied Biosystems, Inc. (Foster City, CA) sequence analyzer.

These results demonstrated that the initial clone (5A-1) contained a 1171 base pair insert (see SEQ ID No. 4). The entire 5A-1 cDNA was found to exist as extended open reading frame for translation into protein. Consequently, it was determined that the 5A-1 cDNA must be a fragment of a larger mRNA.

cDNA Sequence Homologies

Using programs and databases available on the Internet (retrieved from NCBI Blast E-mail Server address blast@ncbi.nlm.nih.gov), it was determined that the 5A-1 clone encodes a previously undefined unique molecule. The BLASTP

program [1.4.8MP, 20-June-1995 (build 11/13/95)] was used to compare all of the possible frames of amino acid sequences encoded by 5A-1 versus all known amino acid sequences available within multiple international databases [Altschul et al., J. Mol. Biol., 215: 403-410 (1990)]. Only one protein, from *Micrococcus luteus*, possessed a marginally significant homology ($p=0.04$) (41%) over a short stretch of 75 of the 390 amino acids encoded by 5A-1. Otherwise, there were not any amino acid homologies (i.e., with $p \leq 0.05$) for any known proteins. Therefore, the protein encoded by 5A-1 is not significantly related to MAO A or B, α_2AR , or any other known eukaryotic protein in the literature.

In contrast to the amino acid search on BLASTP, two nearly homologous EST cDNA sequences of undefined nature covering 155 and 250 b.p. of the 5A-1 clone were reported to exist using BLASTN (reached from the same Internet server on 11/13/95). BLASTN is a program used to compare known DNA sequences from international databases, regardless of whether they encode a polypeptide. Neither of the two EST cDNA sequences having high homology to 5A-1, to our knowledge have been reported anywhere else except on the Internet. Both were derived as Expressed Sequence Tags (ESTs) in random attempts to sequence the human cDNA repertoire [as described in Adams et al., Science, 252: 1651-1656 (1991)]. As far as can be determined, the people who generated these ESTs lack any knowledge of what protein(s) they encode. One cDNA, designated HSA09H122, contained 250 b.p. with 7 unknown/incorrect base pairs (97% homology) versus 5A-1 over

the same region. HSA09H122 was generated in France (Genethon, B.P. 60, 91002 Evry Cedex France) from a human lymphoblast cDNA library. The other EST, designated EST04033, contained 155 b.p. with 12 unknown/incorrect base pairs (92% homology) versus 5A-1 over the same region. EST04033 was generated at the Institute for Genomic Research (Gaithersburg, MD) from a human fetal brain cDNA clone (HFBDP28). Thus, both of these ESTs are short DNA sequences and contain a number of errors (typical of single-stranded sequencing procedures as used when randomly screening ESTs).

Based on the BLASTN search, the owner of HSA09H122 was contacted in an effort to obtain that clone. The current owner of the clone appears to be Dr. Charles Auffret (Paul Brousse Hospital, Genetique, B.P. 8, 94801 Villejuif Cedex, France). Dr. Auffret indicated by telephone that his clone came from a lot of clones believed to be contaminated with yeast DNA, and he did not trust it for release. Contamination with yeast DNA of that clone was later confirmed to have been reported within an Internet database. Thus, HSA09H122 was not reliable.

The other partial clone (EST04033) was purchased from American Type Culture Collection in Rockville, MD (ATCC Catalog no. 82815). A telephone call to the Institute for Genomic Research revealed that it had been deposited at ATCC under [insert terms]. As far as can be determined, the present inventors were the first to completely sequence EST04033. The complete size of EST04033 was 3389 b.p. (SEQ ID No. 1), with a

3,318 b.p. nonplasmid insert (see SEQ ID No. 3). Within this sequence of EST04033 the remaining 783 base pairs of the coding sequence presumed for a 70 kDa imidazoline receptor were predicted at the 5' side of 5A-1 (i.e., 783 coding nucleotides unique to EST04033 + 1171 coding nucleotides of 5A-1 = 1954 predicted total coding nucleotides; assuming b.p.# 1397-1400 in SEQ. No. 1 encodes the initiating methionine). The entire 1954 b.p. coding region for an \approx 70 kDa protein is shown in SEQ ID No. 2. The nucleotide sequence of EST04033 was determined in the same manner as described previously for the 5A-1 clone. The nucleotide sequence of the entire clone is shown in SEQ ID No. 1. In this sequence, an identical overlap was observed for the sequence obtained previously for the 5A-1 clone and the sequence obtained for EST04033. The 5A-1 overlap began at EST04033 b.p. 2,181 (SEQ. No.1) and continued to the end of the molecule (b.p. 3,351).

Conclusions About Our cDNA Clones

cDNA of the present invention encode a protein that is immunoreactive with both of the known selective antisera for an imidazoline receptor, i.e., Reis antiserum and Döntenwill antiserum. Thus, an instant cDNA molecule produces a protein immunologically related to a purified imidazoline receptor and has the antigenic specificity expected for an imidazoline binding site. These antisera have been documented in the scientific literature as being selective for an "imidazoline receptor",

which provides strong evidence that such an imidazoline receptor has indeed been cloned.

As mentioned, our instant cDNA sequence contains open reading frame distinct from any previously described proteins. Therefore, the encoded protein is novel, and it is unrelated to α_2 -adrenoceptors or monoamine oxidases. Small hydrophobic domains in the predicted amino acid sequence suggest that the protein is probably membrane bound, as expected for an imidazoline receptor.

Example 3. Cloning of a Human Gene

A pre-made genomic library of human placental DNA was purchased from Stratagene (La Jolla, CA) to screen for an IR gene by hybridization. The genomic library was constructed in Stratagene's vector λ FIX® II (catalog # 946206), and it was grown in XL1-Blue MRA (P2) host bacteria. It was titered to yield approximately 50,000 plaques per 137 mm plate. Lifts from six such plates were screened in duplicate by hybridization. The DNA probe used for screening was a 1.85 kb EcoR1 fragment from EST 04033 cDNA (uniquely related to our sequences based on the BLASTN). After the restriction digestion of EST 04033 DNA, the 1.85 kb fragment was extracted from an agarose electrophoresis gel, cleaned according to the GENECLEAN® III kit manual (BIO 101, Inc., P.O. Box 2284, La Jolla, CA), and radiolabeled with [α -³²P]d-CTP according to Stratagene's Prime-It® II Random Primer Labeling Kit manual. Plaques were lifted onto

137 mm Duralon-UVTM membranes (Stratagene's catalog #420102), denatured, and cross-linked with Stratagene's UV-StratalinkerTM 1800. Hybridization was conducted under high stringency conditions: prehybridization = 6 X SSC, 1 % SDS, 50 % formamide, and 100 µg/ml of sheared, denatured salmon sperm DNA at 42°C for 2 hrs; hybridization = 6 X SSC, 1 % SDS, 50 % formamide, and 100 µg/ml of sheared, denatured salmon sperm DNA at 45°C overnight; wash = 2 washes of 1 X SSC, 0.1 % SDS at 65°C and 3 washes of 0.2 X SSC, 0.2 % SDS at 65°C. This hybridization procedure is essentially described in Stratagene's vector λ FIX® II instruction manual. Positive plaques were localized by developing Kodak BioMax films. Two positive genomic clones of identical size were retained through three rounds of screening.

One of the positive genomic clones (designated JEP 1-A) was selected for complete characterization. It was found to contain an ≈ 17 kb insert. Large-scale preparations of this genomic clone DNA were performed using the λ QUICK! SPIN kit (BIO101, La Jolla, CA). To verify that we had cloned a gene corresponding to 5A-1 and EST04033 cDNA, some restriction site positions in the genomic clone were determined using the FLASH Nonradioactive Gene Mapping Kit (Stratagene) and compared to Southern blots of human DNA. The location of genomic sequences highly related to (or identical to) those of our cDNA clones was determined by high stringency hybridization (as above) with the following ³²P-labeled probe: a 1110 bp *Apa*I-*Eco*RI fragment from the cDNA clone 5A-1. This fragment was chosen as the probe because it lacks the GAG

repeat (encoding glutamic acids), which might have complicated matters if it were found to be repeated elsewhere in the genome. With genomic clone JEP1-A, we detected a 14.1 kb *EcoRI* fragment and a 7.7 kb *SacI* fragment that hybridized with this probe. Southern blots containing *EcoRI*- or *SacI*-digested human genomic DNA (from human blood) with the 1110 bp *ApaI-EcoRI* cDNA probe also resulted in the detection of a 14.1 kb *EcoRI* fragment and a 7.7 kb *SacI* fragment. No additional restriction fragments of human genomic DNA appeared to hybridize with this probe under lower stringency conditions. These results strongly suggested that this gene (JEP-1A) encodes transcript(s) giving rise to the 5A-1 and EST04033 cDNA clones. Clone JEP-1A has been deposited under the Budapest Treaty with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, MD, USA, 20852, on August 28, 1997 and has been assigned deposit accession no. ATCC 209216.

Genomic DNA sequencing was done by contract with Cadus Pharmaceutical Corporation (Tarrytown, NY). The original lambda JEP1-A clone was subcloned into pZero (Invitrogen) as a convenient vector. The initial fragments for sequencing were derived from *Sac I* and *Xba I* restriction enzymes. The short *Sac I* fragments of 1.5, 3.0 and 3.5 kb were further digested with *Hind III*, *Pst I*, and *Kpn I* yielding 15 subclones of varying length. The procedure consisted of sequencing all these subclones and parent clones with vector forward and reverse primers. Subsequently, this initial round of sequencing was

supplemented with primer walking using custom oligonucleotides. The Sac I fragments were joined together by primer walking using the 2 Xba I fragments of 3 and 10 Kb. Then, the largest Sac I fragment (8 kb) and the 10 kb Xba I fragment were used as templates for a transposon sequencing method. The method used was the Primer Island Transposition Kit (Perkin-Elmer Corp., Norwalk, CT; Applied Biosystems) (ABI). The kit consists of a synthetic transposon (Ty1) containing forward and reverse primers and the integrase enzyme which inserts the transposon randomly into the target plasmid DNA. Transposon insertion is an alternative to subcloning or primer walking when sequencing a large region of DNA (Devine and Boeke, Nucleic Acids Res. 22: 3765-3772 (1994); Devine et al., Genome Res., in press, (1997); Kimmel et al., In Genome Analysis, a Laboratory Manual, Cold Spring Harbor Press, NY, NY, in press (1997)). A total of over 250 individual sequencing reactions were performed. Sequencing was done on ABI model 373 and 377 automated sequencers using ABI dye-terminator sequencing kits. Primers were designed using Gene Runner software (Hastings Software, Hastings On Hudson, NY). Oligonucleotides were purchased from Gibco-BRL (Gaithersburg, MD). Sequence assembly was performed using Sequencer Software (Gene Codes Corp., Ann Arbor, MI) from 4-fold redundancy of sequences.

The entire sequence of our JEP-1A genomic clone is shown in SEQ. 21. The computer program, GENSCAN 1.0, was able to identify splice sites of known topology. As expected, this gene contained

a number of introns. See Table 1 hereinbelow. Only one continuous open reading frame was identified within our genomic clone. This open reading frame was interrupted by a number of introns (which is typical of eukaryotic transcripts) as shown in Fig. 5. The predicted polypeptide is encoded by the genomic DNA beginning at b.p. # 971 of SEQ ID No. 21. The predicted amino acid sequence of the polypeptide encoded thereby is shown in SEQ ID No. 22. As can be seen, the entire 5A-1 DNA and polypeptide sequence was encapsulated within this predicted genomic transcript. Therefore, there is no question that this is the gene encoding 5A-1 and EST04033 cDNA. In addition, JEP-1A has more nearly defined the full-length transcript (by at least 102 more coding nucleotides than the cDNAs alone).

TABLE 1

Position of Predicted Introns and Exons

GENSCAN 1.0 Date run: 26-Aug-97 Time: 12:35:39
Sequence gs_seqfile : 15202 bp : 58.36% C+G : Isochore 4 (57.00 - 100.00 C+G%)
Parameter matrix: HumanIso.smat
Predicted genes/exons:

Gn.Ex	Type	S	.Begin	...End	.Len	Fr	Ph	I/Ac	Do/T	CodRg	P..	Tscr..
1.01	Intr +		971	1084	114	1	0	69	98	200	0.836	20.91
1.02	Intr +		4096	4177	82	0	1	37	53	81	0.358	-0.13
1.03	Intr +		5732	5856	125	0	2	117	95	84	0.953	13.48
1.04	Intr +		6997	7046	50	0	2	95	116	44	0.998	6.52
1.05	Intr +		8416	9825	1410	1	0	96	94	2914	0.970	283.09
1.06	Intr +		10489	10897	409	1	1	15	59	318	0.517	17.19
1.07	Intr +		11293	11449	157	0	1	57	61	236	0.998	18.57
1.08	Intr +		11923	12051	129	2	0	84	63	224	0.997	21.34
1.09	Intr +		12570	12731	162	1	0	95	80	229	0.996	23.94
1.10	Term +		13090	13700	611	2	2	59	41	1012	0.942	89.44
1.11	PlyA +		14257	14262	6							1.05

A BLASTN analysis of the entire genomic sequence (on 08/26/97) demonstrated again that this gene has not been previously defined in the literature.

As with the cDNA clones, some EST sequences of identity were found (listed below and later). Of particular interest was a variance in the first intron splice site predicted by the computer. Upstream of that site (ie., upstream of amino acids PEKKGGG = amino acids predicted after first splice site) we have identified two types of transcripts. Genomic clone JEP-1A predicted 34 amino acids upstream of that sequence before entering another intron upstream. In an identical manner, three ESTs (H61282, AA428790 and AA428250) overlapped that entire region in our clones and they contained the identical nucleotides for those 34 amino acids, plus an additional 22 more amino acids further upstream. By comparison, however, our EST04033 varied from all of these ESTs upstream of that site. This means, the first 1,532 nucleotides of EST04033 (thought to encode translation of amino acids 1-56 of EST04033 beginning at b.p. # 1,398 in SEQ. 1) are completely at variance with the other ESTs down to that splice site, but from there on they are identical. This provides strong evidence that this site can generate two alternatively spliced transcripts which can produce at least one functional protein (ie., the transfections with EST04033). For the reader's information, this splice site is upstream of b.p. # 1,565 in SEQ.1, b.p. # 168 in SEQ.2, b.p. # 1,532 in SEQ.3, amino acid # 57 in SEQ.5, and b.p. # 971 in the genomic SEQ.21.

Genomic Sequence Analysis

Of interest is a unique glutamic- and aspartic acid-rich region within our predicted protein. This region of the IR protein delineates a highly unique span of 59 amino acids, 36 of which are Glu or Asp residues (61%). This region was largely discovered within clone 5A-1 and it is present within all discovered and predicted transcripts from the gene (EST04033 included). This sequence lies between two potential transmembrane loops (hydrophobic domains). The identification of this unique Glu/Asp-rich domain within our clones is consistent with an expected negatively charged pocket capable of binding clonidine and agmatine, both of which are highly positively charged ligands. Also, since the Dontenwill antiserum was specifically developed against an idazoxan/clonidine binding site, and its immunoreactivity is directed against the clone 5A-1/ λ gt11 fusion protein, this suggests that clone 5A-1 might encode an imidazoline binding site. Furthermore, this glu/asp-rich sequence is located within the longest stretch of homology that the clone has with any known protein, i.e., the ryanodine receptor (as determined by on BLASTN). Specifically, we have discovered four regions of homology between the imidazoline receptor and the ryanodine receptor, which are all Glu/Asp-rich. The total nucleic acid homology is 67% with the ryanodine receptor DNA over the stretches encompassing this region. However, this is not sufficient to indicate that the imidazoline receptor is a subtype of the ryanodine receptor, because this

homologous stretch is still a minor portion of the overall transcript(s) identified in the gene. Instead, this significant homology may reflect a commonality in function between this region of the IR and the ryanodine receptor.

The Glu/Asp-rich region within the ryanodine receptor has also been reported to define a calcium and ruthenium red dye binding domain that modulates the ryanodine receptor/ Ca^{++} release channel located within the sarcoplasmic reticulum. The only other charged amino acids within the Glu/Asp-rich region of our clones are two arginines (the ryanodine receptor has uncharged amino acids at the corresponding positions).

Based on this identification of Arg residues within the Glu/Asp-rich region of the predicted imidazoline binding site, the assistance of Dr. Paul Ernsberger (Case Western Reserve University, Cleveland, Ohio) was enlisted. Dr. Ernsberger performed phenylglyoxal attack of arginine on native PC-12 membranes. Dr. Ernsberger was able to demonstrate that this treatment completely eliminated imidazoline binding sites in these membranes. This provides some indirect evidence that the native imidazoline binding site also contains an Arg residue. On the other hand, attempts to chemically modify cysteine and tyrosine residues, which are not located near the Glu/Asp-rich region did not affect PC-12 membrane binding of imidazolines.

As a further test of the sequence, it was determined whether native IR binding sites in PC-12 cells would be sensitive to ruthenium red. From the structure of the cloned sequence, it was

reasoned that native IR should bind ruthenium red. Accordingly, a competition of ruthenium red with ^{125}PIC at native IR sites on PC-12 membranes was studied. In these studies it was observed that ruthenium red competed for ^{125}PIC binding to the same extent as did the potent imidazoline compound, moxonidine, i.e., 100% competition. Furthermore, the IC_{50} for competition of ruthenium red at IR was slightly more robust than reported for ruthenium red on the activation of calcium-dependent cyclic nucleotide phosphodiesterase - the previous most potent effect of ruthenium red on any biological site - indicating possible pharmacological importance. It is also noteworthy that calcium failed to compete for ^{125}PIC binding at PC-12 IR sites (as did a calcium substitute, lanthanum). We and others have previously reported that a number of other cations robustly interfere with IR binding [Ernsberger et al., Annals NY Acad.Sci., 763: 22-42 (1995); Ernsberger et al., Annals NY Acad.Sci., 763: 163-168 (1995)]. Attempts were also made to directly stain the proteins in SDS gels with ruthenium red [Chen and MacLennan, J. Biol. Chem., 269: 22698-22704 (1994)]. It was found that ruthenium red stains the same platelet (33 kDa) and brain (85 kDa) bands that Reis antiserum detects. (Remember, the same 33 kDa band was verified to directly correlate with ^{125}PIC Bmax values for IR.) Thus, these results linked the attributes predicted from the cloned sequence to a native IR binding site.

Two other facets of the predicted polypeptide from JEP-1A suggest that we have identified most of the functional sequences.

First, our predicted protein is comparable in regard to both the order and size of three regions of importance to the function of the interleukin-2R β receptor (IL-2R β). Specifically, IL-2R β possesses the following regions over a span of 286 amino acids: ser-rich region, followed by glu/asp-rich region, followed by proline-rich region. Likewise, our predicted protein has the same three regions, in the same order, over a span of about 625 amino acids. This suggests that our protein might function similarly as cytokine receptors. Secondly, our predicted protein possesses a cytochrome p450 heme-iron ligand signature sequence [Nelson et al., Pharmacogenetics 6: 1-42 (1996)]. This suggests that our protein might also function as do cytochrome p450s in oxidative, peroxidative and reductive metabolism of endogenous compounds.

Some additional findings about the amino acid sequence of our instant IR polypeptide are: (1) The glu/asp-rich region also bears similarity to an amino acid sequence within a GTPase activator protein. (2) There appear to be four small hydrophobic domains indicative of transmembrane domain receptors. (3) A number of potential protein kinase C (PKC) phosphorylation sites appear near to the carboxy side of the protein, and we have previously found that treatment of membranes with PKC leads to an enhancement of native IR binding. Thus, these observations are all consistent with other observations expected for native IR.

RNA Studies

Northern blotting has also been performed on polyA⁺ mRNA from human tissues in order to ascertain the regional expression of the mRNA corresponding to our cDNA. The same 1110 b.p. ApaI-EcoRI fragment from cDNA clone 5A-1 used in Southern blots was used for these studies. This probe region was not found within any other known sequences on the BLASTN database. The results revealed a 6 kb mRNA band, which predominated over a much fainter 9.5 kb mRNA in most regions (Fig. 6). Some exceptions to this pattern were in lymph nodes and cerebellum (Fig. 6), where the 9.5 kb band was equally or more intense. Although the 6 kb band is weakly detectable in some non-CNS tissues, it is enriched in brain. An enrichment of the 6 kb mRNA is observed in brainstem, although not exclusively. The regional distribution of the mRNA is somewhat in keeping with the reported distribution of IR binding sites, when extrapolated across species (Fig. 6). Thus, the rank order of B_{max} values for IR in rat brain has been reported to be frontal cortex > hippocampus > medulla oblongata > cerebellum [Kamisaki et al., Brain Res., 514: 15-21 (1990)]. Therefore, with the exception of human cerebellum, which showed two mRNA bands, the distribution of the mRNA for our the present cloned cDNA is consistent with it belonging to IR.

[It should be noted that while IR binding sites are commonly considered to be low in cerebral cortex compared to brainstem, this is in fact a misinterpretation of the literature based only on comparisons to the alpha-2 adrenoceptor's B_{max}, rather than on absolute values. Thus, IR B_{max} values have actually been

reported to be slightly higher in the cortex than the brainstem, but they only "appear" to be low in the cortex in comparison to the abundance of alpha-2 binding sites in cortex. Therefore, the distribution of the IR mRNA is reasonably in keeping with the actual Bmax values for radioligand binding to the receptor [Kamisaki et al., (1990)].

A final point to emphasize about the Northern blots is that they clearly demonstrate two high-stringency transcripts (Fig. 6). This is in keeping with the alternatively spliced EST cDNAs mentioned earlier. Thus, we suggest this may be the basis for both the 6 and 9.5 kb transcripts.

Summary of Genomic Sequence Results

The JEP-1A clone clearly contains most of the gene. Within it we have identified at least 3,776 nucleotides for transcript(s) (encoding 1,065 amino acids plus 587 b.p. of untranslated region down to the polyT⁺ tail). This has been lengthened by at least 66 coding nucleotides upstream (22 amino acids) in comparison to overlapping ESTs. In addition to this, we are quite confident of the splice site for the two observed mRNA sizes. Most of the functional sequences are predicted to be encoded within our genomic clone.

A summary of the evidence that a gene encoding an imidazoline receptor protein has been cloned is summarized in Table 2 hereinbelow.

TABLE 2

Comparison of Protein Predicted From Our Clones with
Properties of Native IR₁ and I₂ Sites

Imidazoline Receptor-like Clone	Authentic IR ₁	Authentic I ₂
Original λ phage fusion protein (from 5A-1) is immunoreactive with Dontenwill and Reis antibodies	Dontenwill-Ab activity (a) inhibits RVLM IR ₁ binding (³ H-Clonidine), & (b) correlates with 85 kDa Western band. Reis-Ab activity correlates w platelet IR ₁ Bmax (¹²⁵ PIC binding)	Dontenwill & Reis Abs both inhibit brain I ₂ sites (³ H-IDX).
Segment homologous to a GTPase-activator prot'n	Weak to moderate sensitivity to GTP	Not sensitive to GTP
Predicts \geq 120,000 MW protein	85,000 MW immunoreactivity	59-61,000 MW photoaffinity
Predicts 1-4 hydrophobic domains	Enriched in plasma membranes	Enriched in mitochondria
Encodes Glu/Asp-rich (negatively charged) domain consistent with Ca ⁺⁺ and ruthenium red binding	<ul style="list-style-type: none"> • Binds (+)-charged imidazolines • Sensitive to divalent cations • Sensitive to ruthenium red 	<ul style="list-style-type: none"> • Binds (+)-charged imidazolines • Not sensitive to divalent cations • Unknown sensitivity for Ruthen. red
Arginine is only positively charged amino acid near Glu/Asp domain	<ul style="list-style-type: none"> • Arg attack eliminates binding • Cys & Tyr attack w/o effect on binding 	Unknown
Encodes PKC sites	PKC treatment enhances binding	Unknown
Human mRNA Distribution; F.Cortex > hippocampus > medulla	Rat IR ₁ Bmax (¹²⁵ PIC): F.Cortex > hippocampus > medulla	Rat I ₂ Bmax (³ H-IDX): Medulla > F. Cortex
Transfected COS-7 cells expressed high affinity for moxonidine & p-iocloclonidine (PIC)	High affinity for moxonidine and PIC	Low affinity for moxonidine and PIC

Example 4. Transient Transfection Studies

COS-7 cells were transfected with a vector containing EST04033 cDNA, which was predicted based on sequence analysis to contain the glu/asp rich region thought to be important for ligand binding to the imidazoline receptor protein. The EST04033 cDNA was subcloned into pSVK3 (Pharmacia LKB Biotechnology, Piscataway, NJ) using standard techniques [Sambrook, supra], and transfected via the DEAE-dextran technique as previously described [Choudhary et al., Mol.Pharmacol., 42: 627-633 (1992); Choudhary et al., Mol.Pharmacol., 43: 557-561 (1993); Kohen et al., J.Neurochem., 66: 47-56 (1996)]. A restriction map of the EST04033 cDNA is shown in Fig. 3. The restriction enzymes Sal I and Xba I were used for subcloning into pSVK3.

Briefly stated, COS-7 cells were seeded at 3×10^6 cells/100 mm plate, grown overnight and exposed to 2 ml of DEAE-dextran/plasmid mixture. After a 10-15 min. exposure, 20 ml of complete medium (10% fetal calf serum; 5 μ g/ml streptomycin; 100 units/ml penicillin, high glucose, Dulbeccos' modified Eagle's medium) containing 80 μ M chloroquine was added and the incubation continued for 2.5 hr. at 37°C in a 5% CO₂ incubator. The mixture was then aspirated and 10 ml of complete medium containing 10% dimethyl sulfoxide was added with shaking for 150 seconds.

Following aspiration, 15 ml of complete medium with dialyzed serum was added and the incubation continued for an additional 65 hours. After this time period, the cells from 6 plates were harvested and membranes were prepared as previously described

[Ernsberger et al., Annals NY Acad. Sci., 763: 22-42 (1995), the disclosure of which is incorporated herein by reference].

Parent, untransfected COS-7 cells were prepared as a negative control. Some membranes were treated with and without PKC for 2 hrs prior to analysis, since previous studies had indicated that receptor phosphorylation could be beneficial to detect IR binding.

Transfected samples were also analyzed by Western blots.

The protocol used for Western blot assay of transfected cells is as follows. Cell membranes were prepared in a special cocktail of protease inhibitors (1 mM EDTA, 0.1 mM EGTA, 1 mM phenylmethyl-sulfonylfluoride, 10 mM ϵ -aminocaproic acid, 0.1 mM benzamide, 0.1 mM benzamide-HCl, 0.1 mM phenanthroline, 10 μ g/ml pepstatin A, 5 mM iodoacetamide, 10 μ g/ml antipain, 10 μ g/ml trypsin-chymotrypsin inhibitor, 10 μ g/ml leupeptin, and 1.67 μ g/ml calpain inhibitor) in 0.25 M sucrose, 1 mM $MgCl_2$, 5 mM Tris, pH 7.4. Fifteen μ g of total protein were denatured and separated by SDS gel electrophoresis. Gels were equilibrated and electrotransferred to nitrocellulose membranes. Blots were then blocked with 10% milk in Tris-buffered saline with 0.1% Tween-20 (TBST) during 60 min. of gentle rocking. Afterwards, blots were incubated in anti-imidazoline receptor antiserum (1:3000 dil.) for 2 hours. Following the primary antibody, blots were washed and incubated with horseradish peroxidase-conjugated anti-rabbit goat IgG (1:3000 dil.) for 1 hr. Blots were extensively washed and incubated for 1 min. in a 1:1 mix of Amersham ECL detection

solution. The blots were wrapped in cling-film (SARAN WRAP) and exposed to Hyperfilm-ECL (Amersham) for 2 minutes. Quantitation was based on densitometry using a standard curve of known amounts of protein containing BAC membranes or platelet membranes run in each gel.

One nM [125 I]p-iodoclonidine was employed in the radioligand binding competition assays, since at this low concentration this radioligand is selective for the IR site much more than for I_2 binding sites. The critical processes of membrane preparation of tissue culture cells and the radioligand binding assays of IR and I_2 have been reviewed by Piletz and colleagues [Ernsberger et al., Annals NY Acad Sci., 763: 510-519 (1995)]. Total binding ($n = 12$ per experiment) was determined in the absence of added competitive ligands and nonspecific binding was determined in the presence of 10^{-4} M moxonidine ($n = 6$ per experiment). Log normal competition curves were generated against unlabeled moxonidine, p-iodoclonidine, and (-) epinephrine. Each concentration of the competitors was determined in triplicate and the experiment was repeated thrice.

The protocol to fully characterize radioligand binding in the transfected cells entails the following. First, the presence of IR and/or I_2 binding sites are scanned over a range of protein concentrations using a single concentration of [125 I]-p-iodoclonidine (1.0nM) and 3 H-idazoxan (8nM), respectively. Then, rate of association binding experiments (under a 10μ M mask of NE to remove α_2 AR interference) are performed to determine if the

kinetic parameters are similar to those reported for native imidazoline receptors [Ernsberger et al. Annals NY Acad. Sci., 763: 163-168 (1995)]. Then, full Scatchard plots of [125 I]-p-iodoclonidine (2-20 nM if like IR) and 3 H-idazoxan (5-60 nM if like I_2) binding are conducted under a 10 μ M mask of NE. Total NE (10 μ M)-displaceable binding is ascertained as a control to rule out α_2 -adrenergic binding. The Bmax and K_D parameters for the transfected cells are ascertained by computer modeling using the LIGAND program [McPherson, G., J.Pharmacol.Meth., 14: 213-228 (1985)] using 20 μ M moxonidine to define IR nonspecific binding, or 20 μ M cirazoline to define I_2 nonspecific binding.

The results of the transient transfection experiments of the imidazoline receptor vector into COS-7 cells are shown in Fig. 4. Competition binding experiments were performed using membrane preparations from these cells and 125 PIC was used to radiolabel IR sites. A mask of 10 μ M norepinephrine was used to rule out any possible α_2 AR binding in each assay even though parent COS-7 cells lacked any α_2 AR sites. Moxonidine and p-iodoclonidine (PIC) were the compounds tested for their affinity to the membranes of transfected cells. As can be seen, the affinities of these compounds in competition with 125 PIC were well within the high affinity (nM) range.

The following IC_{50} values and Hill slopes were obtained in this study: moxonidine, IC_{50} = 45.1 nM (Hill slope = 0.35 ± 0.04); p-iodoclonidine without PKC pretreatment of the membranes, IC_{50} = 2.3 nM (Hill slope = 0.42 ± 0.06); p-iodoclonidine with PKC

pretreatment of the membranes, $IC_{50} = 19.0 \text{ nM}$ (Hill slope = 0.48 ± 0.08). Shallow Hill slopes for [^{125}I]p-iodoclonidine have been reported before in studies of the interaction of moxonidine and p-iodoclonidine with the human platelet IR_1 binding site [Piletz and Sletten, (1993)]. Epinephrine failed to displace any of the [^{125}I]p-iodoclonidine binding in the transfected cells, as expected since this is a nonadrenergic imidazoline receptor. Furthermore, in untransfected cells less than 5% of the amount of displaceable binding was observed as for the transfected cells - and this "noise" in the parent cells all appeared to be low affinity (data not shown). These results thus demonstrate the high affinities of two imidazoline compounds, p-iodoclonidine and moxonidine, for the portion of our cloned receptor encoded within EST04033. PKC pretreatment of the membranes had no effect in the transfected COS cells.

It was also observed that the level of the expressed protein, as measured by Western blotting of the transfected cells, was consistent with the level of IR binding that was detected. In other words, a protein band was uniquely detected in the transfected cells, and it was of a density consistent with the amount of radioligand binding. Hence, the present results are in keeping with those expected for an imidazoline receptor. In summary, these data provide direct evidence that the EST04033 clone encodes an imidazoline binding site having high affinities for moxonidine and p-iodoclonidine, which is expected for an IR protein.

Example 5. Stable Transfection Methods.

Stable transfections can be obtained by subcloning the imidazoline receptor cDNA into a suitable expression vector, e.g., pRC/CMV (Invitrogen, San Diego, CA), which can then be used to transform host cells, e.g. CHO and HEK-293 cells, using the Lipofectin reagent (Gibco/BRL, Gaithersburg, MD) according to the manufacturer's instructions. These two host cell lines can be used to increase the permanence of expression of an instant clone. The inventors have previously ascertained that parent CHO cells lack both alpha₂-adrenoceptor and IR binding sites [Piletz et al., J. Pharm. & Exper. Ther., 272: 581-587 (1995)], making them useful for these studies. Twenty-four hours after transfection, cells are split into culture dishes and grown in the presence of 600 µg/ml G418-supplemented complete medium (Gibco/BRL). The medium is changed every 3 days and clones surviving in G418 are isolated and expanded for further investigation.

Example 6. Direct Cloning of More Complete Gene and Other Homologous Human IR.

Direct probing of other human genomic and cDNA libraries can be performed by preparing labelled cDNA probes from different subcloned regions of our clone. Commercially available human DNA libraries can be used. Besides the cDNA and genomic libraries we have already screened, another genomic library is EMBL (Clontech), which integrates genomic fragments up to 22 kbp long.

It is reasonable to expect that introns may exist within other human IR genes so that only by obtaining overlapping clones can the full-length genes be sequenced. A probe encompassing the 5' end of an instant cDNA is generally useful to obtain the gene promoter region. Clontech's Human PromoterFinder DNA Walking procedure provides a method for "walking" upstream or downstream from cloned sequences such as cDNAs into adjacent genomic DNA.

Example 7. Methods for Preparing Antibodies to Imidazoline Receptive Proteins.

An instant imidazoline receptive polypeptide can also be used to prepare antibodies immunoreactive therewith. Thus, synthetic peptides (based on deduced amino acid sequences from the DNA) can be generated and used as immunogens. Additionally, transfected cell lines or other manipulations of the DNA sequence of an instant imidazoline receptor can provide a source of purified imidazoline receptor peptides in sufficient quantities for immunization, which can lead to a source of selective antibodies having potential commercial value.

In addition, various kits for assaying imidazoline receptors can be developed that include either such antibodies or the purified imidazoline receptor protein. A purification protocol has already been published for the bovine imidazoline receptor in BAC cells [Wang et al, 1992] and an immunization protocol has also been published [Wang et al., 1993]. These same protocols can be utilized with little if any modification to afford

purified human IR protein from transfected cells and to yield selective antibodies thereto.

In order to obtain antibodies to a subject peptide, the peptide may be linked to a suitable soluble carrier to which antibodies are unlikely to be encountered in human serum. Illustrative carriers include bovine serum albumin, keyhole limpet hemocyanin, and the like. The conjugated peptide is injected into a mouse, or other suitable animal, where an immune response is elicited. Monoclonal antibodies can be obtained from hybridomas formed by fusing spleen cells harvested from the animal and myeloma cells [see, e.g., Kohler and Milstein, Nature, 256: 495-497 (1975)].

Once an antibody is prepared (either polyclonal or monoclonal), procedures are well established in the literature, using other proteins, to develop either RIA or ELISA assays [see, e.g., "Radioimmunoassay of Gut Regulatory Peptides; Methods in Laboratory Medicine," Vol. 2, chapters 1 and 2, Praeger Scientific Press, 1982]. In the case of RIA, the purified protein can also be radiolabelled and used as a radioactive antigen tracer.

Currently available methods to assay imidazoline receptors are unsuitable for routine clinical use, and therefore the development of an assay kit in this manner could have significant market appeal. Suitable assay techniques can employ polyclonal or monoclonal antibodies, as has been previously described [U.S. Patent No. 4,376,110 (issued to David et al.), the disclosure of

which is incorporated herein by reference].

Summary

In summary, we have identified unique DNA sequences that have properties expected of a gene and the cDNA transcript(s) of an imidazoline receptor. Prior to our first cloning the cDNA, only two sequences of EST cDNA were identified within public databases having similar nature. But, these were both partial and imprecise sequences - not identified at all with respect to any encoded protein. Indeed, one of them (HSA09H122) was reported to be contaminated. In our hands, the other EST 04033 clone was correctly sequenced for the first time (in its entirety = 3318 bp). Prior to this, even the size of EST 04033 was unknown. The present inventors also demonstrated that an imidazoline receptive site can be expressed in cells transfected with the EST 04033 cDNA clone, and this site has the proper potencies of an IR. We have deduced most of the complete cDNA encoding this protein.

The present invention has been described with reference to specific examples for purposes of clarity and explanation. Certain obvious modifications of the invention readily apparent to one skilled in the art can be practiced within the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: John E. Piletz

Tina R. Ivanov

(ii) TITLE OF INVENTION: DNA Sequence Encoding a Human
Imidazoline Receptor

(iii) NUMBER OF SEQUENCES: 22

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: WENDEROTH, LIND & PONACK

(B) STREET: 805 Fifteenth St. N.W., Suite 700

(C) CITY: Washington

(D) STATE: District of Columbia

(E) COUNTRY: U.S.A.

(F) ZIP: 20005

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.50" 1.44 Mb diskette

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: MS-DOS

(D) SOFTWARE: Wordperfect 5.1+

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: new

(B) FILING DATE:

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NO.: 60/012,600

(B) FILING DATE: March 1, 1996

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NO.: 08/650,766

(B) FILING DATE: May 20, 1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Warren Cheek

(B) REGISTRATION NUMBER: 33,367

(C) REFERENCE/DOCKET NUMBER: WMC-1342/clone

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (202) 371-8850

(B) TELEFAX: (202) 371-8856

(2) INFORMATION FOR SEQ ID NO: 1

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3389 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA

(B) CLONE: EST04033 (HFBDP28)

(ix) FEATURE:

(A) NAME/KEY: predicted translation product when
transfected

(B) LOCATION: 1398 ... 3389

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1

CTCTAGAAC TAGTGGATCC CCCGGGCTGC AGGAATTCCA GTTTAATACT AACCCATAATG	60
GTGACTGCG GTTTACAAAG AGCTCTGTAT CACCTGGGAT AGCTTTCAGT AGCAATTCAC	120
ACAACTGGT CCTAAAAAAT AATAACAATA ATAATAATAA TTAGAGAATT AAAACCCAAC	180
SCATGTTGA ATGGTTAAAA TCACGTAAGA ACTGAAATTT GGGGTGGGGG TGTCTCTAAC	240
ECTGAGCTT GTCCTAGCAG TGAANAATGCT CGCCTCCAAG CAGGGCTCAG AAAGGTCTGG	300
CCCCTCCAG GCAGAGGGCT GAGCTCAGGG GGCTCTTGA GGACACTCAC CCCATGGTCC	360
FGGATGCT TCTGGCTTCC TTAANAACAG TTGGGCATCC GCATTGTATA AGTAGGTGA	420
ACCCTAGTG TGGTCTTTTT GAAGGATATG GGAAGGGAGG ATGACGAACT AGAGAAGTGG	480
AGGGGACCA AAATCACTGA GGTCCCAGAA TATCATAGAT TTGGGTATAG GATTGGGGTC	540
TAAGAATT GAGCACCAGG AATTCCAGCT TCTTCCCATT AAAGAACTG GGACTGGTTT	600
ECCTTGGAG GCCTATGTAG TGTTTTCTGC CCCTGTCCA TACCAAGTCT CATTGATATT	660
ETGCAGAAT ATCAGATGAA AATCTATTTT TAAAGACCAT TGGGAGAATG GGTGTGGGAG	720
AGGAGTTGG AGTGGGGTTG GGGGGCAGTT AAAAATGAAT AAAAATCTCT CAGCTACAGA	780
CCCAAACAT CACTTCCCTC CGCATTACAC GCATTTCCCA GCAGTCCCCA GATGGTTGTT	840
CCCTGGGGA CACAGCAGCT GCCTCATTTT CCTTCAGGCC CCATGGGCTG CTGGTCAACC	900
AGGATCTA CTAAAGATGA CGCAAATGCC GACTGAACAA TCTGAAACCC AAAGGACTCG	960
AGAGAGACA TGTTCGTCTG AGGAGAGAAA GGTGAGCCAA GGGCAGGGCC CAGGTCCCCC	1020
GGGGCCCC CGAGAGCCCC GACATGCACC TTCTGGATGT GTTGTGTTCAA GTAGGACTTA	1080
AGCGGAAGA AGCTCCACAC TTCAGGGCAT GGTACTTCT TCTCCCCATC AGACTCCATT	1140
GTTTTTTG GGACTGCCAT GTCGAGGAG AAAGAGCCAT TGGCACTCTG CTCTCTGGC	1200
CTTCAGGT CGCTGGCATC TGAGAGGTCA CCATAGGAGT CAGAGCTCTC AATCGGATCC	1260
ATGTGAGC ATTTCTGGCC TTCTCGGTTA CAGATACTGC AGAAGTTGCT GGGCCCCCTG	1320
GTGCTTCT TCAGGTGGTC TGCCATGTAT GCTGCCCGCA AGTACTTCCC ACACACCTGG	1380
GGGCACCT TGTCTTC ATG ACA GGC CAG GTG GGA GCG CAG ACG GTC TCG	1430

Met Thr Gly Gln Val Gly Ala Gln Thr Val Ser

1

5

10

1478
 15 20 25
 1478
 15 20 25

1526
 30 35 40
 1526
 30 35 40

1574
 45 50 55
 1574
 45 50 55

1622
 65 70 75
 1622
 65 70 75

1670
 80 85 90
 1670
 80 85 90

1718
 95 100 105
 1718
 95 100 105

CC AGC AGC CTC TCG TCC ACT GAC AGT CTG ACT CCC GAG CAC CAG CCC 1766
 la Ser Ser Leu Ser Ser Thr Asp Ser Leu Thr Pro Glu His Gln Pro
 110 115 120

TT GCC CAG GGA TGT TCT GAT TCC TTG GAG TCC ATC CCT GCG GGA CAG 1814
 le Ala Gln Gly Cys Ser Asp Ser Leu Glu Ser Ile Pro Ala Gly Gln
 125 130 135

CA GCT TCC GAT GAT TTA AGG GAC GTG CCA GGA GCT GTT GGT GGT GCA 1862
 la Ala Ser Asp Asp Leu Arg Asp Val Pro Gly Ala Val Gly Gly Ala
 40 145 150 155

GC=CCA GAA CAT GCC GAG CCG GAG GTC CAG GTG GTG CCG GGG TCT GGC 1910
 er Pro Glu His Ala Glu Pro Glu Val Gln Val Val Pro Gly Ser Gly
 160 165 170

AG=ATC ATC TTC CTG CCC TTC ACC TGC ATT GGC TAC ACG GCC ACC AAT 1958
 ln Ile Ile Phe Leu Pro Phe Thr Cys Ile Gly Tyr Thr Ala Thr Asn
 175 180 185

AG GAC TTC ATC CAG CGC CTG AGC ACA CTG ATC CGG CAG GCC ATC GAG 2006
 ln Asp Phe Ile Gln Arg Leu Ser Thr Leu Ile Trp Gln Ala Ile Glu
 190 195 200

3G CAG CTG CCT GCC TGG ATC GAG GCT GCC AAC CAG CGG GAG GAG GGC 2054
 rp Gln Leu Pro Ala Trp Ile Glu Ala Ala Asn Gln Trp Glu Glu Gly

205

210

215

AG GGT GAA CAG GGC GAG GAG GAG GAT GAG GAG GAG GAA GAA GAG GAG 2102
 In Gly Glu Gln Gly Glu Glu Glu Asp Glu Glu Glu Glu Glu Glu
 20 225 230 235

AC GTG GCT GAG AAC CGC TAC TTT GAA ATG GGG CCC CCA GAC GTG GAG 2150
 sp Val Ala Glu Asn Arg Tyr Phe Glu Met Gly Pro Pro Asp Val Glu
 240 245 250

AG GAG GAG GGA GGA GGC CAG GGG GAG GAA GAG GAG GAG GAA GAG GAG 2198
 lu Glu Glu Gly Gly Gly Gln Gly Glu Glu Glu Glu Glu Glu Glu
 255 260 265

AT GAA GAG GCC GAG GAG GAG CGC CTG GCT CTG GAA TGG GCC CTG GGC 2246
 sp Glu Glu Ala Glu Glu Glu Arg Leu Ala Leu Glu Trp Ala Leu Gly
 270 275 280

CG GAC GAG GAC TTC CTG CTG GAG CAC ATC CGC ATC CTC AAG GTG CTG 2294
 la Asp Glu Asp Phe Leu Leu Glu His Ile Arg Ile Leu Lys Val Leu
 285 290 295

CG TGC TTC CTG ATC CAT GTG CAG GGC AGT ATC CGC CAG TTC GCC GCC 2342
 rp Cys Phe Leu Ile His Val Gln Gly Ser Ile Arg Gln Phe Ala Ala
 300 305 310 315

GC CTT GTG CTC ACC GAC TTC GGC ATC GCA GTC TTC GAG ATC CCG CAC	2390
ys Leu Val Leu Thr Asp Phe Gly Ile Ala Val Phe Glu Ile Pro His	
320 325 330	

AG GAG TCT CGG GGC AGC AGC CAG CAC ATC CTC TCC TCC CTG CGC TTT	2438
ln Glu Ser Trp Gly Ser Ser Gln His Ile Leu Ser Ser Leu Arg Phe	
335 340 345	

TC TTT TGC TTC CCG CAT GGC GAC CTC ACC GAG TTT GGC TTC CTC ATG	2486
al Phe Cys Phe Pro His Gly Asp Leu Thr Glu Phe Gly Phe Leu Met	
350 355 360	

CG GAG CTG TGT CTG GTG CTC AAG GTA CGG CAC AGT GAG AAC ACG CTC	2534
ro Glu Leu Cys Leu Val Leu Lys Val Arg His Ser Glu Asn Thr Leu	
365 370 375	

TC ATT ATC TCG GAC GCC GCC AAC CTG CAC GAG TTC CAC GCG GAC CTG	2582
he Ile Ile Ser Asp Ala Ala Asn Leu His Glu Phe His Ala Asp Leu	
80 385 390 395	

GC TCA TGC TTT GCA CCC CAG CAC ATG GCC ATG CTG TGT AGC CCC ATC	2630
rg Ser Cys Phe Ala Pro Gln His Met Ala Met Leu Cys Ser Pro Ile	
400 405 410	

TC TAC GGC AGC CAC ACC AGC CTG CAG GAG TTC CTG CGC CAG CTG CTC	2678
eu Tyr Gly Ser His Thr Ser Leu Gln Glu Phe Leu Arg Gln Leu Leu	

415

420

425

CC TTC TAC AAG GTG GCT GGC GGC TGC CAG GAG CGC AGC CAG GGC TGC 2726
 ar Phe Tyr Lys Val Ala Gly Gly Cys Gln Glu Arg Ser Gln Gly Cys
 430 435 440

FC CCC GTC TAC CTG GTC TAC AGT GAC AAG CGC ATG GTG CAG ACG GCC 2774
 ae Pro Val Tyr Leu Val Tyr Ser Asp Lys Arg Met Val Gln Thr Ala
 445 450 455

CC GGG GAC TAC TCA GGC AAC ATC GAG TGG GCC AGC TGC ACA CTC TGT 2822
 la Gly Asp Tyr Ser Gly Asn Ile Glu Trp Ala Ser Cys Thr Leu Cys
 50 465 470 475

TA GCC GTG CGG CGC TCC TGC TGC GCG CCC TCT GAG GCC GTC AAG TCC 2870
 ar Ala Val Arg Arg Ser Cys Cys Ala Pro Ser Glu Ala Val Lys Ser
 480 485 490

CC GCC ATC CCC TAC TGG CTG TTG CTC ACG CCC CAG CAC CTC AAC GTC 2918
 la Ala Ile Pro Tyr Trp Leu Leu Leu Thr Pro Gln His Leu Asn Val
 495 500 505

CC AAG GCC GAC TTC AAC CCC ATG CCC AAC CGT GGC ACC CAC AAC TGT 2966
 e Lys Ala Asp Phe Asn Pro Met Pro Asn Arg Gly Thr His Asn Cys
 510 515 520

:C AAC CGC AAC AGC TTC AAG CTC AGC CGT GTG CCG CTC TCC ACC GTG 3014
 :g Asn Arg Asn Ser PHe Lys Leu Ser Arg Val Pro Leu Ser Thr Val
 525 530 535

:G CTG GAC CCC ACA CGC AGC TGT ACC CAG CCT CGG GGC GCC TTT GCT 3062
 :u Leu Asp Pro Thr Arg Ser Cys Thr Gln Pro Arg Gly Ala Phe Ala
 :O 545 550 555

:T GGC CAC GTG CTA GAG CTG CTC GTG GGG TAC CGC TTT GTC ACT GCC 3110
 :p Gly His Val Leu Glu Leu Leu Val Gly Tyr Arg Phe Val Thr Ala
 560 565 570

:C TTC GTG CTG CCC CAC GAG AAG TTC CAC TTC CTG CGC GTC TAC AAC 3158
 :e Phe Val Leu Pro His Glu Lys Phe His Phe Leu Arg Val Tyr Asn
 575 580 585

:G CTG CGG GCC TCG CTG CAG GAC CTG AAG ACT GTG GTC ATC GCC AAG 3206
 :n Leu Arg Ala Ser Leu Gln Asp Leu Lys Thr Val Val Ile Ala Lys
 590 595 600

:C CCC GGG ACG GGA GGC AGC CCC CAG GGC TCC TTT GCG GAT GGC CAG 3254
 :r Pro Gly Thr Gly Gly Ser Pro Gln Gly Ser Phe Ala Asp Gly Gln
 605 610 615

:T GCC GAG CGC AGG GCC AGC AAT GAC CAG CGT CCC CAG GAG GTC CCA 3302
 :o Ala Glu Arg Arg Ala Ser Asn Asp Gln Arg Pro Gln Glu Val Pro

20

625

630

635

CA GAG GCT CTG GCC CCG GCC CCA GTG GAA GTC CCA GCT CCA GCC CCG 3350

la Glu Ala Leu Ala Pro Ala Pro Val Glu Val Pro Ala Pro Ala Pro

640

645

650

AA TTC GAT ATC AAG CTT ATC GAT ACC GTC GAC CTG CAG

3389

lu Phe Asp Ile Lys Leu Ile Asp Thr Val Asp Leu Gln

655

660

664

2) INFORMATION FOR SEQ ID NO: 2

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1954 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 2

TCACAGGCC AGGTGGGAGC GCAGACGGTC TCGGGTGGCA AAAGAAGCAT TGCAGGTCTG 60

CAGTTGTGA GGCCGCTCAG AAGTGTGCAC CTGCTTGATA TGTCCGTTCA AGTGATCAGG 120

CTGGAGAAG CCTTTCCAC AGCTCTGGCA GATGTAAGGC GGAATTCCCC AGAGAAGAAG 180

STGGTGAAG ACTCCCGGCT CTCAGCTGCC CCCTGCATCA GACCCAGCAG CTCCCCTCCC 240

CTGTGGGCTC CCGCATCTGC CTCCTTGCCC CAGCCCATCC TCTCTAACCA AGGAATCATG 300

FCGTTACAG AGGAGGCCCT GGCCAGCAGC CTCTCGTCCA CTGACAGTCT GACTCCCGAG 360

ACCAGCCCA TTGCCCAGGG ATGTTCTGAT TCCTTGGAGT CCATCCCTGC GGGACAGGCA 420

TTTCCGATG ATTTAAGGGA CGTGCCAGGA GCTGTTGGTG GTGCAAGCCC AGAACATGCC 480

AGCCGGGAGG TCCAGGTGGT GCCGGGGTCT GGCCAGATCA TCTTCCTGCC CTTACCTGCG 540

TTGGCTACA	CGGCCACCAA	TCAGGACTTC	ATCCAGCGCC	TGAGCACACT	GATCCGGCAG	600
CCATCGAGC	GGCAGCTGCC	TGCCTGGATC	GAGGCTGCCA	ACCAGCGGGA	GGAGGGCCAG	660
GTGAACAGG	GCGAGGAGGA	GGATGAGGAG	GAGGAAGAAG	AGGAGGACGT	GGCTGAGAAC	720
GCTACTTTG	AAATGGGGCC	CCCAGACGTG	GAGGAGGAGG	AGGGAGGAGG	CCAGGGGGAG	780
AAGAGGAGG	AGGAAGAGGA	GGATGAAGAG	GCCGAGGAGG	AGCGCCTGGC	TCTGGAATGG	840
CCCTGGGCG	CGGACGAGGA	CTTCCTGCTG	GAGCACATCC	GCATCCTCAA	GGTGTGTGG	900
GCTTCCTGA	TCCATGTGCA	GGGCAGTATC	CGCCAGTTCG	CCGCTGCCTT	TGTGCTCACC	960
ACTTCGGCA	TCGCACTCTT	CGAGATCCCG	CACCAGGAGT	CTCGGGGCAG	CAGCCAGCAC	1020
TCCTCTCTT	CCCTGCGCTT	TGTCTTTTGC	TTCCCGCATG	GCGACCTCAC	CGAGTTTGGC	1080
TCCTCATGC	CGGAGCTGTG	TCTGGTGCTC	AAGGTACGGC	ACAGTGAGAA	CACGCTCTTC	1140
TTATCTCGG	ACGCCGCCAA	CCTGCACGAG	TTCCACGCGG	ACCTGCGCTC	ATGCTTTGCA	1200
CCGAGCACA	TGGCCATGCT	GTGTAGCCCC	ATCCTCTACG	GCAGCCACAC	CAGCCTGCAG	1260
AGTTTCTGC	GCCAGTGCTT	CACCTTCTAC	AAGGTGGCTG	GCGGCTGCCA	GGAGCGCAGC	1320
AGGCTGCTT	TCCCCGTCTA	CCTGGTCTAC	AGTGACAAGC	GCATGGTGCA	GACGGCCGCC	1380
GGGACTACT	CAGGCAACAT	CGAGTGGGCC	AGCTGCACAC	TCTGTTCAAG	CGTGCGGGCG	1440
CCGCTGCGG	CGCCCTCTGA	GGCCGTCAAG	TCCGCCGCCA	TCCCTACTAG	GCTGTTGCTC	1500
CGCCCCAGC	ACCTCAACGT	CATCAAGGCC	GACTTCAACC	CCATGCCCAA	CCGTGGCACC	1560
ACAACTGTC	GCAACCGCAA	CAGCTTCAAG	CTCAGCCGTG	TGCCGCTCTC	CACCGTGCTG	1620
TGGACCCCA	CACGCAGCTG	TACCCAGCCT	CGGGGCGCCT	TTGCTGATGG	CCACGTGCTA	1680
AGCTGCTCG	TGGGGTACCG	CTTGTCTACT	GCCATCTTCG	TGCTGCCCCA	CGAGAAGTTC	1740
ACTTCTGTC	GCGTCTACAA	CCAGCTGCGG	GCCTCGCTGC	AGGACCTGAA	GACTGTGGTC	1800
TCGCCAAGA	CCCCCGGGAC	GGGAGGCAGC	CCCCAGGGCT	CCTTTGCGGA	TGGCCAGCCT	1860
CCGAGCGCA	GGGCCAGCAA	TGACCAGCGT	CCCCAGGAGG	TCCCAGCAGA	GGCTCTGGCC	1920
CGGCCCCAG	TGGAAGTCCC	AGCTCCAGCC	CCGG			1954

3) INFORMATION FOR SEQ ID NO: 3

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3318 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 3

ATTCCAGTT TAATACTAAC CCTAATGTGT GACTGCGGTT TACAAAGAGC TCTGTATCAC	60
TGGGATAGC TTTCAGTAGC AATTCAC TAC AACTGGTCCT AAAAAATAAT AACAAATAATA	120
TAATAATTA GAGAATTAAA ACCCAACAGC ATGTTGAATG GTTAAATCA CGTAAGAACT	180
AAATTGGG GTGGGGGTGT CCTCAACAGC TGAGCTTGTC CTAGCAGTGA AAATGCTCGC	240
TCGAAGCAG GGCTCAGAAA GGTCTGGAGC CCTCCAGGCA GAGGGCTGAG CTCAGGGGGC	300
CTTGGAGGA CACTCACCCC ATGGTCCATG GGATGCTTCT GGCTTCCTTA AAAACAGTTG	360
GCATCCGCA TTGTATAAGT AGGTGGAGAC CCTAGTGTGG TTCTTTTGAA GGATATGGGA	420
GGCAGGATG ACGAACTAGA GAAGTGGGAG GGGACCAAAA TCACTGAGGT CCCAGAATAT	480
ATAGATTG GGTATAGGAT TGGGCTCACT AAGAATTGAG CACCAGGAAT TCCAGCTTCT	540
CCGATTAAA GAAACTGGGA CTGGTTTTCG CTTGGAGGCC TATGTAGTGT TTTCTGCCCC	600
GTGCCATAC CAAGTCTCAT TGATATTCT CAGAATATC AGATGAAAAT CTATTTCTAA	660
GATCATTGG GAGAATGGGT GGTGGAGAAG GACTTGAGT GGGGTTGGGG GGCAGTTAAA	720
ATGAATAAA AATCTCTCAG CTACAGAAC CAAACATCAC TTCCCTCCGC ATTCACAGCA	780
TTCCCAGCA GTCCCCAGAT GGTGTTTTC GTGGGGACAC AGCAGCTGCC TCATTTCCCT	840
CAGGCCCCA TGGGCTGCTG GTCAACCTCA GGATCTACTA AAGATGACGC AAATGCCGAC	900
GAACAATCT GAAACCCAAA GGAATCGAGG AGAGACATGT TCTGCTGAGG AGAGAAAGGT	960
AGCCAAGGG CAGGGCCGAG GTCCCCCAGG GGGCCCCCGA GAGCCCGGAC ATGCACCTTC	1020
GGATGTGTT TGTTCAAGTA GGAATAGAG CGGAAGAAGC TCCACATTC AGGGCATGGG	1080
ACTTCTTCT CCCCATCAGA CTCCATTTTG TTTTGGGGA CTGCCATGTC GCAGGAGAAA	1140

AGCCATTGG	CACTCTGCTT	CTCTGGCGTC	TTCAGGTCGC	TGGCATCTGA	GAGGTCACCA	1200
AGGAGTCAG	AGCTCTCAAT	CGGATCCTGA	TGTGAGCATT	TCTGGCCTTC	TCGGTTACAG	1260
FACTGCAGA	AGTTGCTGGG	CCCCTCGCTG	TGCTTCTTCA	GGTGGTCTGC	CATGTATGCT	1320
CCGCAAGT	ACTTCCCACA	CACCTGGCAG	GGCACCTTGT	CTTCATGACA	GGCCAGGTGG	1380
AGCGCAGAC	GGTCTCGGGT	GGCAAAAGAA	GCATTGCAGG	TCTGACACTT	GTGAGGCCCG	1440
DAGAAGTGT	GCACCTGCTT	GATATGTCCG	TTCAAGTGAT	CAGGCCTGGA	GAAGCCTTTC	1500
DACAGCTCT	GGCAGATGTA	AGGCGGAATT	CCCCAGAGAA	GAAGGGTGGT	GAAGACTCCC	1560
ECTCTCAGC	TGCCCCCTGC	ATCAGACCCA	GCAGCTCCCC	TCCCAGTGTG	GCTCCCGCAT	1620
FGCTCCCT	GCCCCAGCCC	ATCCTCTCTA	ACCAAGGAAT	CATGTTCTGT	CAGGAGGAGG	1680
CTGGCCAG	CAGCCTCTCG	TCCACTGACA	GTCTGACTCC	CGAGCACCAG	CCCATTGCCC	1740
GGGATGTT	TGATTCCCTT	GAGTCCATCC	CTGCGGGACA	GGCAGCTTCC	GATGATTTAA	1800
GGAGCTGCC	AGGAGCTGTT	GGTGGTGCAA	GCCCAGAACA	TGCCGAGCCG	GAGGTCCAGG	1860
GGTGGCCGG	GTCTGGCCAG	ATCATCTTCC	TGCCCTTCAC	CTGCATTGGC	TACACGGCCA	1920
GAATCAGGA	CTTCATCCAG	CGCCTGAGCA	CACTGATCCG	GCAGGCCATC	GAGCGGCAGC	1980
CCCTGCCTG	GATCGAGGCT	GCCAACCAGC	GGGAGGAGGG	CCAGGGTGAA	CAGGGCGAGG	2040
GGAGGATGA	GGAGGAGGAA	GAAGAGGAGG	ACGTGGCTGA	GAACCGCTAC	TTTGAAATGG	2100
CGCCCCAGA	CGTGGAGGAG	GAGGAGGGAG	GAGGCCAGGG	GGAGGAAGAG	GAGGAGGAAG	2160
GGAGGATGA	AGAGGCCGAG	GAGGAGCGCC	TGGCTCTGGA	ATGGGCCCTG	GGCGCGGACG	2220
GGACTTCCT	GCTGGAGCAC	ATCCGCATCC	TCAAGTGCT	GTGGTGCTTC	CTGATCCATG	2280
CAGGGCAG	TATCCGCCAG	TTCGCCCGCT	GCCTTGCTGT	CACCGACTTC	GGCATCGCAG	2340
TTTCAGAT	CCCGCACCAG	GAGTCTCGGG	GCAGCAGCCA	GCACATCCTC	TCCTCCCTGC	2400
TTTGTCTT	TTGCTTCCCG	CATGGCGACC	TCACCGAGTT	TGGCTTCCTC	ATGCCGGAGC	2460
TGTCTGGT	GCTCAAGGTA	CGGCACAGTG	AGAACACGCT	CTTCATTATC	TCGGACGCCG	2520
AACCTGCA	CGAGTTCCAC	GCGGACCTGC	GCTCATGCTT	TGCACCCAG	CACATGGCCA	2580
CTGTGTAG	CCCCATCCTC	TACGGCAGCC	ACACCAGCCT	GCAGGAGTTC	CTGCGCCAGC	2640
CTCACCTT	CTACAAGGTG	GCTGGCGGCT	GCCAGGAGCG	CAGCCAGGGC	TGCTTCCCCG	2700

TACCTGGT CTACAGTGAC AAGCGCATGG TGCAGACGGC CGCCGGGGAC TACTCAGGCA	2760
ATCGAGTG GGCCAGCTGC ACACTCTGTT CAGCCGTGCG GCGCTCCTGC TGC GCGCCCT	2820
GAGGCCGT CAAGTCCGCC GCCATCCCCT ACTGGCTGTT GCTCAGCCCC CAGCACCTCA	2880
GTCACTAA GGCCGACTTC AACCCCATGC CCAACCGTGG CACCCACAAC TGTCGCAACC	2940
AACAGCTT CAAGCTCAGC CGTGTGCCGC TCTCCACCGT GCTGTGGAC CCCACACGCA	3000
TGTACCA GCCTCGGGGC GCCTTTGCTG ATGGCCACGT GCTAGAGCTG CTCGTGGGT	3060
CGCTTTGT CACTGCCATC TTCGTGCTGC CCCACGAGAA GTTCCACTTC CTGCGCGTCT	3120
AACCAGCT GCGGGCCTCG CTGCAGGACC TGAAGACTGT GGTCACTGCC AAGACCCCGG	3180
ACGGGAGG CAGCCCCCAG GGCTCCTTTG CGGATGGCCA GCCTGCCGAG CGCAGGGCCA	3240
AATGACCA GCGTCCCCAG GAGGTCCCAG CAGAGGCTCT GGCCCCGGCC CCAGTGAAG	3300
CCAGCTCC AGCCCCGG	3318

INFORMATION FOR SEQ ID NO: 4

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1171 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 4

3GAGGAGG AAGAGGAGGA TGAAGAGGCC GAGGAGGAGC GCCTGGCTCT GGAATGGGCC	60
3GGCGCGG ACGAGGACTT CCTGCTGGAG CACATCCGCA TCCTCAAGGT GCTGTGGTGC	120
CTGATCC ATGTGCAGGG CAGTATCCGC CAGTTCGCCG CTTGCCTTGT GCTCACCAGC	180
GGGATCG CAGTCTTCGA GATCCCGCAC CAGGAGTCTC GGGGCAGCAG CCAGCACATC	240
TCCTCCC TGCGCTTTGT CTTTGTCTC CCGCATGGCG ACCTCACCGA GTTTGGCTTC	300
TATGCCGG AGCTGTGTCT GGTGCTCAAG GTACGGCACA GTGAGAACAC GCTCTTCATT	360

CTCGGACG CCGCCAACCT GCACGAGTTC CACGCGGACC TGCGCTCATG CTTTGCACCC	420
GCACATGG CCATGCTGTG TAGCCCCATC CTCTACGGCA GCCACACCAG CCTGCAGGAG	480
CCTGCGCC AGCTGCTCAC CTTCTACAAG GTGGCTGGCG GCTGCCAGGA GCGCAGCCAG	540
CTGCTTCC CCGTCTACCT GGTCTACAGT GACAAGCGCA TGGTGACAGC GGCCGCCGGG	600
CTACTCAG GCAACATCGA GTGGGCCAGC TGCACACTCT GTTCAGCCGT GCGGCGCTCC	660
CTGCGCGC CCTCTGAGGC CGTCAAGTCC GCCGCCATCC CTTACTGGCT GTTGCTCAGC	720
CCAGCACC TCAACGTCAT CAAGGCCGAC TTCAACCCCA TGCCCAACCG TGGCACCAC	780
CTGTCGCA ACCGCAACAG CTTCAAGCTC AGCCGTGTGC CGCTCTCCAC CGTGCTGCTG	840
CCCCACAC GCAGCTGTAC CCAGCCTCGG GCGCCTTTG CTGATGCECA CGTGCTAGAG	900
GCTCGTGG GGTACCGCTT TGTCACTGCC ATCTTCGTGC TGCCCCACGA GAAGTTCAC	960
CCTGCGCG TCTACAACCA GCTGCGGGCC TCGCTGCAGG ACCTGAAGAC TGTGGTCATC	1020
CAAGACCC CCGGGACGGG AGGCAGCCCC CAGGGCTCCT TTGCGGATGG CCAGCCTGCC	1080
GCCSAGGG CCAGCAATGA CCAGCGTCCC CAGGAGGTCC CAGCAGAGGC TCTGGCCCCG	1140
CCCAGTGG AAGTCCCAGC TCCAGCCCCG G	1171

(5) INFORMATION FOR SEQ ID NO: 5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 651 amino acids
- (B) TYPE: polypeptide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 5

Met Thr Gly Gln Val Gly Ala Gln Thr Val Ser
1 5 10

Gly Gly Lys Arg Ser Ile Ala Gly Leu Thr Leu Val Arg Pro Leu Arg
15 20 25

Ser Val His Leu Leu Asp Met Ser Val Gln Val Ile Arg Pro Gly Glu
30 35 40

Ala Phe Pro Thr Ala Leu Ala Asp Val Arg Trp Asn Ser Pro Glu Lys
45 50 55

Lys Gly Gly Glu Asp Ser Trp Leu Ser Ala Ala Pro Cys Ile Arg Pro
60 65 70 75

Ser Ser Ser Pro Pro Thr Val Ala Pro Ala Ser Ala Ser Leu Pro Gln
80 85 90

Pro Ile Leu Ser Asn Gln Gly Ile Met Phe Val Gln Glu Glu Ala Leu
95 100 105

Ala Ser Ser Leu Ser Ser Thr Asp Ser Leu Thr Pro Glu His Gln Pro
110 115 120

Ile Ala Gln Gly Cys Ser Asp Ser Leu Glu Ser Ile Pro Ala Gly Gln
125 130 135

Ala Ala Ser Asp Asp Leu Arg Asp Val Pro Gly Ala Val Gly Gly Ala
140 145 150 155

Ser Pro Glu His Ala Glu Pro Glu Val Gln Val Val Pro Gly Ser Gly
160 165 170

Gln Ile Ile Phe Leu Pro Phe Thr Cys Ile Gly Tyr Thr Ala Thr Asn
175 180 185

Gln Asp Phe Ile Gln Arg Leu Ser Thr Leu Ile Trp Gln Ala Ile Glu
190 195 200

Trp Gln Leu Pro Ala Trp Ile Glu Ala Ala Asn Gln Trp Glu Glu Gly
205 210 215

Gln Gly Glu Gln Gly Glu Glu Glu Asp Glu Glu Glu Glu Glu Glu
220 225 230 235

Asp Val Ala Glu Asn Arg Tyr Phe Glu Met Gly Pro Pro Asp Val Glu
240 245 250

Glu Glu Glu Gly Gly Gly Gln Gly Glu Glu Glu Glu Glu Glu
255 260 265

Asp Glu Glu Ala Glu Glu Glu Arg Leu Ala Leu Glu Trp Ala Leu Gly
270 275 280

Ala Asp Glu Asp Phe Leu Leu Glu His Ile Arg Ile Leu Lys Val Leu
285 290 295

Trp Cys Phe Leu Ile His Val Gln Gly Ser Ile Arg Gln Phe Ala Ala
300 305 310 315

Cys Leu Val Leu Thr Asp Phe Gly Ile Ala Val Phe Glu Ile Pro His
320 325 330

Gln Glu Ser Trp Gly Ser Ser Gln His Ile Leu Ser Ser Leu Arg Phe
335 340 345

Val Phe Cys Phe Pro His Gly Asp Leu Thr Glu Phe Gly Phe Leu Met
350 355 360

Pro Glu Leu Cys Leu Val Leu Lys Val Arg His Ser Glu Asn Thr Leu
365 370 375

Phe Ile Ile Ser Asp Ala Ala Asn Leu His Glu Phe His Ala Asp Leu
380 385 390 395

Arg Ser Cys Phe Ala Pro Gln His Met Ala Met Leu Cys Ser Pro Ile
400 405 410

Leu Tyr Gly Ser His Thr Ser Leu Gln Glu Phe Leu Arg Gln Leu Leu
415 420 425

Thr Phe Tyr Lys Val Ala Gly Gly Cys Gln Glu Arg Ser Gln Gly Cys
430 435 440

Phe Pro Val Tyr Leu Val Tyr Ser Asp Lys Arg Met Val Gln Thr Ala
445 450 455

Ala Gly Asp Tyr Ser Gly Asn Ile Glu Trp Ala Ser Cys Thr Leu Cys
460 465 470 475

Ser Ala Val Arg Arg Ser Cys Cys Ala Pro Ser Glu Ala Val Lys Ser
480 485 490

Ala Ala Ile Pro Tyr Trp Leu Leu Leu Thr Pro Gln His Leu Asn Val
495 500 505

Ile Lys Ala Asp Phe Asn Pro Met Pro Asn Arg Gly Thr His Asn Cys
510 515 520

Arg Asn Arg Asn Ser Phe Lys Leu Ser Arg Val Pro Leu Ser Thr Val
525 530 535

Leu Leu Asp Pro Thr Arg Ser Cys Thr Gln Pro Arg Gly Ala Phe Ala
540 545 550 555

Asp Gly His Val Leu Glu Leu Leu Val Gly Tyr Arg Phe Val Thr Ala
560 565 570

Ile Phe Val Leu Pro His Glu Lys Phe His Phe Leu Arg Val Tyr Asn
575 580 585

Gln Leu Arg Ala Ser Leu Gln Asp Leu Lys Thr Val Val Ile Ala Lys
590 595 600

Thr Pro Gly Thr Gly Gly Ser Pro Gln Gly Ser Phe Ala Asp Gly Gln
605 610 615

Pro Ala Glu Arg Arg Ala Ser Asn Asp Gln Arg Pro Gln Glu Val Pro
620 625 630 635

Ala Glu Ala Leu Ala Pro Ala Pro Val Glu Val Pro Ala Pro Ala Pro
640 645 650

(6) INFORMATION FOR SEQ ID NO: 6

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 390 amino acids
- (B) TYPE: polypeptide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 6

Glu Glu Glu Glu Glu
1 5

Asp Glu Glu Ala Glu Glu Glu Arg Leu Ala Leu Glu Trp Ala Leu Gly
10 15 20

Ala Asp Glu Asp Phe Leu Leu Glu His Ile Arg Ile Leu Lys Val Leu
25 30 35

Trp Cys Phe Leu Ile His Val Gln Gly Ser Ile Arg Gln Phe Ala Ala
40 45 50

Cys Leu Val Leu Thr Asp Phe Gly Ile Ala Val Phe Glu Ile Pro His
55 60 65 70

Gln Glu Ser Trp Gly Ser Ser Gln His Ile Leu Ser Ser Leu Arg Phe
75 80 85

Val Phe Cys Phe Pro His Gly Asp Leu Thr Glu Phe Gly Phe Leu Met
90 95 100

Pro Glu Leu Cys Leu Val Leu Lys Val Arg His Ser Glu Asn Thr Leu
105 110 115

Phe Ile Ile Ser Asp Ala Ala Asn Leu His Glu Phe His Ala Asp Leu
120 125 130

Arg Ser Cys Phe Ala Pro Gln His Met Ala Met Leu Cys Ser Pro Ile
135 140 145 150

Leu Tyr Gly Ser His Thr Ser Leu Gln Glu Phe Leu Arg Gln Leu Leu

155

160

165

Thr Phe Tyr Lys Val Ala Gly Gly Cys Gln Glu Arg Ser Gln Gly Cys
170 175 180

Phe Pro Val Tyr Leu Val Tyr Ser Asp Lys Arg Met Val Gln Thr Ala
185 190 195

Ala Gly Asp Tyr Ser Gly Asn Ile Glu Trp Ala Ser Cys Thr Leu Cys
200 205 210

Ser Ala Val Arg Arg Ser Cys Cys Ala Pro Ser Glu Ala Val Lys Ser
215 220 225 230

Ala Ala Ile Pro Tyr Trp Leu Leu Leu Thr Pro Gln His Leu Asn Val
235 240 245

Ile Lys Ala Asp Phe Asn Pro Met Pro Asn Arg Gly Thr His Asn Cys
250 255 260

Arg Asn Arg Asn Ser Phe Lys Leu Ser Arg Val Pro Leu Ser Thr Val
265 270 275

Leu Leu Asp Pro Thr Arg Ser Cys Thr Gln Pro Arg Gly Ala Phe Ala
280 285 290

Asp Gly His Val Leu Glu Leu Leu Val Gly Tyr Arg Phe Val Thr Ala
295 300 305 310

Ile Phe Val Leu Pro His Glu Lys Phe His Phe Leu Arg Val Tyr Asn
315 320 325

Gln Leu Arg Ala Ser Leu Gln Asp Leu Lys Thr Val Val Ile Ala Lys
330 335 340

Thr Pro Gly Thr Gly Gly Ser Pro Gln Gly Ser Phe Ala Asp Gly Gln
345 350 355

Pro Ala Glu Arg Arg Ala Ser Asn Asp Gln Arg Pro Gln Glu Val Pro

Ala Glu Ala Leu Ala Pro Ala Pro Val Glu Val Pro Ala Pro Ala Pro
375 380 385 390

(7) INFORMATION FOR SEQ ID NO: 7

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 7

CTTGAGGATG CCGATGTGCT 20

(8) INFORMATION FOR SEQ ID NO: 8

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 8

CCATGGGGTG AGTGTCTT 18

(9) INFORMATION FOR SEQ ID NO: 9

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 9

AGGACACTCA CCCCATGG 18

(10) INFORMATION FOR SEQ ID NO: 10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 10

GTATGGGACA GGGGCAGAAA 20

(11) INFORMATION FOR SEQ ID NO: 11

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 11

TTTCTAAAGA CCATTGGGAG 20

(12) INFORMATION FOR SEQ ID NO: 12

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 12

CCATTTTAAA GTAGCGGTTG 20

(13) INFORMATION FOR SEQ ID NO: 13

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 13

AGGAGAGAAA GGTGAGCCAA 20

(14) INFORMATION FOR SEQ ID NO: 14

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 14

GTAGATCCTG AGGTTGACCA 20

(15) INFORMATION FOR SEQ ID NO: 15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 15

TGTGAGCATT TCTGGCCTTC 20

(16) INFORMATION FOR SEQ ID NO: 16

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 16

TGAAGACGCC AGAGAAGCAG 20

(17) INFORMATION FOR SEQ ID NO: 17

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 17

GCCTCACAAG TGTCAGACCT 20

(18) INFORMATION FOR SEQ ID NO: 18

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 18

AGAAGGGTGG TGAAGACT 18

(19) INFORMATION FOR SEQ ID NO: 19

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 19

CTTGTTAGA GAGGATGGGC 20

(20) INFORMATION FOR SEQ ID NO: 20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 20

(21) INFORMATION FOR SEQ ID NO: 21

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15202 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY:
- (B) LOCATION:
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION: /note="N is unknown or other"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21

GATCCGAGCTCAATTAACCTCACTAAAGGGAGTCGACTCGATCCTTAAA
 ATATTATATCTCCTGGACAACCTGTGGCCATAGTGCCTGACTGTAAACC
 CAAAGGGTTTGCCTTTGCCAGTGTAGCCCAGCCTGGTGTCTGCTGCCCTT
 CGCGGTGTCTGTGCACCTGCCACGATGCTGACCAGACACCTTAACCAGG
 TTCACCCATCGCCTGGGCTGGAGCAGTCCCCCTGATGCTCTGATTGGTC
 CTTGACCTTCTGTCTCCCAAAATCCCAGGTGAGAAATACCTGGAAGT
 CTATTGTGTCCACCTCCCTCTTTGTGGCCGCAAGTGCCCCCTTCTCTCA
 CACAGTCACAAGACCATGAGATGCCATCTCCTCCCTCCTGGGCTGCAGA
 CTTTGGGAAGCTCCCAGGCCACAGAGGTGTGAGCTCCTGTCCAGGCCCTT
 GGGACCTTCCCTCATTCAACCACCTACCCAACCCCCCACTGCCTGCCAG
 CCACCACTCCCTCCCACATTGTCAGGCGGGGGCCCTGCCCTCTCCTGCCG
 CTGGTTCCCTACCCAGGAGGCTCTCCCATCGCTCTTTTGAGAGTCTGCC

TCCACCTCTAACTGGGGCTTAGTTCAAGTTGCCCCCTTACCCTAGTCC
CAGCTGCCCAAGAGCTTGCTGCCTCCTGTTCTTGGTGAGGGACTCCAGAG
ACAGATGTGAGACCTCCCTGGACCCCTCCAAGGCATTCCCAGGTCACTTC
CATGAGTAGTGAGAACCGCCTCTGAGCAGGCTGAGCCTCCCTCAGCCTA
TGGTGTCTCACGTGGCTTGGGCCACAGCAGGTGCTCAGGCCTCCTCCTC
AGCAGAGCCTACCATCCTCCTGCCATGCTCACCAGTCCCCATGCTGATAG
CCATCACCAGTCCCCATGCTGATAGCCATCACCAGTCCCCATGCTGATAG
CCACTTCTGGATGCTCTAGGTCTGTCTGGATGACACAGTGACCACAGAG
AAGGAGCTGGACACTGTGGAAGTGCTGAAAGCAATTCAGAAAGCCAAGGA
GGTCAAGTCCAAACTGAGCAACCCAGAGAAGAAGGTGGGTTTGTGTGGCA
GGTGGGAGGGCAGTGGTGCAGAGCCAGCCGGATAGGAGCCAGTTCGGGG
GGCTTGGGCCATGGGACTGCTCAGGGCTGCCGAGTCCCAGCTGCGCCCCT
CCCTGGCTGCATGACCTCGGGCAAGTCGCGCCCTCTCTGTTCTCTGTGGG
GTGGGGACAGTGGTAGTTCCCTGCTCTAAGGATATGATGAGACCATCTTTA
CCACCCAGTTGGTGGGAACCGTTGCGCTCCCTCCTCACACCCTGGCCCTT
GGGGAGCTCTGTGCTTCCTCTCTCTCCCGGGCTGACTCAAGCACTCGTC
CTCAGGGTGGTGAAGACTCCCGGCTCTCAGCTGCCCCCTGCATCAGACCC
AGCAGCTCCCTCCCACTGTGGCTCCCGCATCTGCCTCCCTGCCCCAGCC
CATCCTCTCTAACCAAGGTAATCGTGTATGTATCTTGCTTCTAGTGGAGC
CACACAGCCCTGCCTGGGCCCCCTGGCTGGGCTGGGGTTGGGGGAGAGGT
GCCAGCACCTGCTTCCAACAGGGTCAGACACAGGGAGGGCAGTGCCTTCT
GCAGGCTGGTCTCGCGGGGGACACATGGCAGGGGTGCCTGGCCTGATG
CCAGCTGTTGCTTGCTTGGTGAGGACTCCCAATTGCTCTGATGCCACAT
CCAGCTCCTCTAGGAGACCGCAGGGTGTCTGACAGGCCCTGAGGCTGCCC
TCTGAACAGGCTCGGGGCTGTTGGCTCATGGGACCCATTCCCTCACCGGC
AGCACAAGCAGGTTGGCTCCTGGTTACAGGAAGCCGGGCTTGTGACTTTA

CTGTCTGGAGCCCGAATCCCTGTGCAGGGAAAAGCTTGCTTTTATCACTG
 CCTCATCTCTGTGGGGTGACCCAGCCCCAGAACACCATGTTTGTGGGGCC
 AAGATGGGCCATCTCTGTCCCTGTGGACCCATGGAAGACCAGGCCCATTC
 GTCTGCCCACTATCTTAGCGTTTTCAAAGGGCTTTACCTCTGAACCCAG
 GCATCCTCGGAGATGAGTGAGTGAAGCAGGTCTCATGAGCGTGTCTGCTG
 GCGCGCCCCCACGGAAGAGGGGAGGGTGTGCCGTCCCGAGTGGAGCCGA
 GGCTCGGGACACGCAGGAAAGGACGCCCGCTGCCCGGGCTCCTGGAGACG
 CAGAACTTGGTGTGAGGTCTTGGGAAAAAGTTCAACCCGATGTTTTAAG
 AGCCAGAAAAACATTCCCACCCCTTGACCTGGTAACCCCACTGGTGGGA
 TTTTCTCTTAGAGGGATAAGATACCGGGAAGGGGAGGTGAAATGCTCACC
 ACTGCCAAAACACGGGCTGCAACTGCAACATCGGAGGATGAGAGGGAGAG
 TCGGCTGTGGTGCAGAATGCTCAGCAGCCCTCCCAGCAGGGACAGGAAGA
 CTGGGCAGGAAGAGGGGAGAAGCATTCAAGTTAAGGCAAAAGGCCCAACG
 CAGAGCAGCACACTGAGGTACACCTGTGAGATGTGGAAGAGAATTCTGT
 AGCGTGGAGCGATGGGGTTAGGTGCCAGGATGATTGCCCATTTTGCTTCT
 GTCAGACTCTTGACTAAGGATTCTGTGTTGCATTTATTACATAAAAGCC
 AGGGAGGTTATATCACGGTGAGAAAGCTTCCCTGACGCCGCCTCCTGTAG
 CGCAGCCAAGCGAGCCTGTGGAGGTACCATATGACTGTAGGCCCTCTGGGG
 ACAGGGAGCTGCATCTGCTTCTCAAGGCCAGGGACACAGCCATTTCTGCC
 AGCATCTGTTGATCAGTGAGTGAGTGAGTGGGCAGGTAGAGCAGGAGCCA
 GTGAAGAGCAGGCCCTGGATGGGTGGGGATGCACCATGTCCCCAGGCTGC
 AGCTGCAGGCAGCCCCCACATTGTGCGAGAAGCCTCTGCACCAGCTCAG
 CCCCCCTCTCACTCCCCCTGTGCCCTGGGGACACTCTGCAGAGGGGCACT
 CTGCAGTCTGTCCCCGCCATCGTGGACTTCTGGACATGGCCTCCAGATT
 TGCACCTCTTAAATAAATCTGCAGTGGATGTCTTTGTGTGCACCTCTCTT
 TCCTTTTGGTGAGAAACAGCAAAGATCGGACCCCTAAGGACTCTCCTGAT

GTCTCCGCTCTATCCGCTGAGTGCCCTTTCTGACCACTGTTTGTACAGG
CCACGGTCCAGGACGGGAGCAGATAGACTGTCCCTGTCCCTGTCCACATT
TCCTTGGTCCAAACAGGGCTTGTGGGAGGTAGTGGCAAAGGTGTTGGTC
TTTTTCTCACTGATTGGAGGCCCTCCCGTGTGTTTTTTCAGCCGCGTGT
TCCTGGGTCTTGCCTGGATGGACAGGGTTTTTTAGCGGTGGGAGCAGCT
TTGCTGACCATGCTGTGCTTCCAGCCTGATCCCGAGAAGGGAGCGTG
CTTCGGAAGGAACTGGCACTCGGGCCTGCCTGAAGGGGGCGCTGTCCAGA
CACACCCAGCCTCCCGTCGTGGCAGGCGCTGTCGGAGCCATGGATGATTG
TGACCAATAGGGGTGGTCGCCAGAGTTGATTGTCCAGCCAGGCCAGGGG
CTGAGAGGAGGCTGTGTGGAGAGGTGGTTAGGAGCCAGGGGCTCGGTACAG
TGAGTTCGCATGCCAGCTTCCTAGCTGTGGGACCTCAAGCAACTTGTAGC
CCCTCTGAAGCTGTTTTCTCAACTGTGAAGTGGACGCACCCTACTTCATT
GATTCTAAGAGGCACGCATTTCCACCTTGTGACTTCTCTGAAACTGAGGT
GCGTCTTTCACTCAGTCAGTGGCGCTCTCATAGTCGCTGTCAGCCAGCTGGTATT
CGAGATGGAGTCGTGGAACCCGTGGACACCTTCCGCTAGGACCAAGAT
GGCGCCACCTGCCGCATCTTAGATTTGATGAAATGTGGTAAATAACGAGA
GGCATGCATGAGCGAATGCTGGGGAGGCGCTTGGCACTACCCAGAGCTCC
ACAGAGGTGGTCGATGAGGGCTGCCCTTTCCACATCCTTAGTAGGGGGT
TCAAGATGACCCAGACTGTGCCCTGGGGAGCTTGGAGCCATGCGGGAGG
ATGAGCCATGTGCTGGAGGAGAACAGGGTAGGATGGTGTGGGGCTTTTGT
AGACTGTCTAGAGCAGAGAAGGTCTGCAGTGGAGGTGGTGTCTGAGGTGA
ATCTCGAAGGTGAATAGGAGTTGAACGTTAGCAGGCAGAGGGTGGATTGC
AGGAGAGCAGCGGCTGGGCAGGTGCCACGCTGGCCCATCAGGGTGCTT
CATGCATGGCTGTGTGCTTGCCATCCTTCCTGCCCTGCCCTACCCCTGCTG
CTTCGCTTCATGGGGCGTTTGAGCTTGGGGCCACCTGCCCTGCCCTCGCTT
GTGGGCAGAGGACCCAGGCTGTGTGAGTTGTCTGTCCCGGGGAGCAGCT

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

GAGCTTGTCCGGGTTCTCTCGACCTGTGGGGCTTCAGAGGACTTCGGGTCA
TTTCAATGGGCTGTGGCGATGCTGGCTGTGGAGGTAGCCTAGGGCTCCTG
TAGCCTTCAGTGAGACTGGCGGCCCCGATGCCAGTGTTACCCCTGCTGGC
GGCAGTCAGGAACATGTTCAAAAGCTTTACTTCAAGTGCTTAGAGGTG
ATCTGAGGTGGAGTAACAGGTCCAGATAGGCTACGTTCAATAAACAGCTT
CAGCGGGGTTTAGGAACACTGTGCATTTACGGGACGCAGTGGGTGAGAGT
GCTGCTGTCCGTGGGAGGTGGCCCCAGGGCAGGTCAGTGGGCACGTCCTG
TGGTAAGTGGGACTGTGGATGTGGGCTCAGGCTGGACTCAGCAGCCCTGC
TGGATACCAAGGCCTGCAAGGGCTGGCCCCCTGGTGAATTGTCCCGTGCC
CTGTGTATCTATGAGTCCTGCAGAGATGACAAATCAGGGGACGGGGTCAT
GTCTAGTCACCGTCTGGGAAAATGCTCCAGGAGTGAACACATTTCAGGCT
CTTGATGGATGTACCTCCAAACTCTTCTCTGGATGGGTGGGCCAGCTTGC
ATGCCTGTGCCGGCCTCTGCCCAGCGAGGTCAGGGCCAGGCCACACAGTC
AGTCTGACTTTGGCAGAAAGTTGAGAGGCAACACTTGTCTCTTGTTTTCAGC
TTGCCTTTCTTTGTGTACTTCTGAGAGCGAGCATTTCTTTTCATGTTCTAT
CCGCTGGCCGTTCTTCTGCGGAATGTCTGTTACGTCCTTTGCACTCTGT
TAATGAGGTTTCCAACCTTCCCTCATTTTTGTAACTCTGTAAGAACTTTTT
CCAGACTAGCGATATAAATCCTGTCAAATATTGCAAACACTTTTCTCAT
TTCATCTGGTTTTAATCTATCCTGGTTTTTAAAAAATGTGTCTGTGGAAG
TTTAATTTTTATGTAGTCACATCTCAGTTTTTTTCCATTGCAATTTATTCT
CAGAATGCTTCTCCCTGCCCTGAGATTAGATAAGCAGTCATTGTGTTCTTT
CTTGAGTTATTTTGTAGATTTTCACTTTTAACATTTTCTTCTATAATCCATG
TGGCTGGGTTTTGGGATCTGGCTAACCCCGCCATGCCAGTAGCCTGAGG
GGCCAGCCCCACTTGTGTAACAGCCGCTCTCCCGCCCCACCCACCCTG
CCTGCCTGCCACCCGCCCTGCTCTCCAGGAATCATGTTTCGTTTCAGGA
GGAGGCCCTGGCCAGCAGCCTCTCGTCCACTGACAGTCTGACTCCCGAGC

ACCAGCCCATGCCCAGGGATGTTCTGATTCTCTGGAGTCCATCCCCTGCG
GGACAGGTAATGCCCTCTTCCCGCTTCTGGGGACCATACATCTGTGGGTG
GACTCTTCTGCTTGGGGTTGTGTGCAGTAGGAAGTGGCCTAGCTGGAGCT
GAGGCAGATGCTTCCAGGGTTTGGCGTCCTCTGCTTTGCGCCACGGTCTT
TCTCTTGGACCTGTCTCTGGTTGAGTGTCTTCTGACAAACACAGTGGTT
AAGGTTTATTTTCAGCCTCCCTCCTTCCCTTCCCCACCCACCTTGTTG
ATGGGAACAGGCAGTTCTCTGTCACTGGGCCCAGGGCACGAGGGGGCAG
GTGGAGAGGGTGGCCCTTGACCTGTGAGCAGGCTTCCCTGGGGAAGGCA
TTTCAAAGACCCTCGTGCAGGGGCTTGTTTGGGTTTCTTCTCTGTTTC
TGGCACCCCTGGAGCCACTCGGCGCCTTTCGCGATGTACCCCTGGTGGTC
TGGGAAACAGTCTCACTCTGGCGCCTCCTCTGTGGTTGTTACTGAGAGTT
CTGGGGCCCCCTTCTTTGTCTGAGGAAAGACAGAGGAAAGCAAGGGTG
CTTGCTGTGTGCTTCGCAAATGTGCTTGGTGCCTGGGCCTCCCTCCAGCC
CCATCTCTGCAGCAGCACAAAGTTATGGCCTTGTGACACTGGGACAGTTT
GCAGAGTCTTGTCTGTCTCCTAGTACTCCACAGTATTCTGCCATCACCC
TTCCAGGGTCACACAGCAAGAGATTCCCAAGCCCTAGGTATTCCCCAGTG
CACAGAGACCATTGGGAGGGACTTGCCAGGGCTGTGTCCACTGCTGGCCA
GTTAGGGTCGGACCAAATTTGTAGACTGTCTACCTGGACCCTTGCCTGGC
ACAAGGAGCAGTCAGATGCTGGATCCCTGGAGAGTGGCGAGAGGCTCTGG
CCTTAGGTTGCGAGTGGGAATCCCAGCCCTGCTGTGTGCTGGTGGGATAA
CCAAGTGGGTCTCTGCCCTTGGGTCCAGAGTGGGCCCCAGGGTCCCAGA
GTGGGCTCCAGGGTACAGCGTGGGGATGGGGAGCCTCCTCAGGGCGGTGA
TGGAGGGCAGAAATGCCAGCTCAGGGTCTGGCAACCAGTAAATGGCTGGG
GCTGGCTGCAGTAGTGGGGACTGACTGTGTTTCTTTCTCCATCAGGCAG
CTTCCGATGATTTAAGGGACGTGCCAGGAGCTGTTGGTGGTGCAAGGTAA
GGAAGAGGTTGGAAAGGGACCTGGGCCTGGCCACACAGCCTTATGCACAC

A C A C T G C T G T G G G C C A G G G G T G G C C A G T C A G G T T T T T T A A A A A T C C G T T
 C A C A G A A G G C C T A T A G A A C T A T T T C T T C C T C T A A A G A G A C A C A G A T G A G A
 T G G A C T T T T C A A T C T G T T T C C A A A T T C T A A T A C C T A A A C T C T G C T C A G C A
 C A T G T T G C C C T A C A C C A G G G G T G G C A A T C A A G G C C T G T G T G T G G C C C A
 C A G C C T G G G A G C T A A G A A T G A C A G T T A C A T T C T T T T T C T T T T T T G A G A
 C T G A G T C T C G C T C T G T C G C C C A G G C T G G A G T G C A G T G G C G T T T C T T G G C
 T C A C T G C A A C C C C C C C T C C C A G A T T A A T G C A A T T T T C C T G T C T C A G C C T
 C A G C C T T C T G A G T A G C C C G A C C A C A G G C G C A C G C C A C C A C G C C C A A C T A
 A T T T T T T A T A T T T T T A G T A G A G A C A G A G A T T C A C C A T G T G G C C T A G C T G G
 T C T C G A A C T C C T G A A C T C C A G T A T C C A C C A C C T C G G C T T C C T A A A G T A
 C T G G A A T T A C A G G C A T G A G C C A C C G C G C C T G G C T A G A A T A C A G T T A C T T
 T T T T T T C T T T G A G A C T G A G T C T T G C T T T G T C A C C C A G G C T G G A G T G C A G
 T G G C A C G A T C T C A G C T C G C T G C A A C C T C C G C C T C C G G G T T C A A G C G A T T
 C T T C T G C C T C A G C C A C C A A G G T G C C C G C C A C C A C A C C T G G C T A A T T T T T
 C T G T T T T T A G T A G G G A C A G G A T T T C G C C A T G T T G G A C A G T T A C A T T C T T A
 A A G G G C T G C T G A A G A T C G T A T G G A C A T G G T A G C C C A T A A A T C C C A A A A T G
 T G T A C T C T G A C C C T T T A C A G A A G C T T A C T A A C T C C C A C T C T A C A T G T G A G
 G G C T G C G G T G G C C A A G A A G A G C T G G A A T T T A A G T G T G A A G G T C C T A A G A C
 C T G C C C C A G C C C A C T T C C C T G C C C C G A G G C C A C C A G G G G T G A C A A G T A G
 A T T C A T G C C C T G G A G T G T T C C T T C T C T C G G G G C T T A T G G C A G C A A C T G A
 A T G A C T T A G A A G T C C A T G G G A G T G C T T T C T G T T G T G G G A A C T C G T G T G G T
 C T G G G C A T A G C T G T G C C A G G C A C C T A T G G T C C A A G C C C C T A G A A G C A T A G
 A C T C T G A C C A A A C T G G C G A C C C A G C C T T C C A G C A G G C A G C A C T G G C T C C C
 A C C A G G G C C C T C A T C C T G G G A A C T G A C T T G G C C A T G T G G G A G G C T T G G G A
 G A C C C A T G G G T T G G T T T C T C A G G G T C A G G G T G T A G C A G T G G G C T C C A G A T
 G T G G C A G G T G G G A G G T G G G A G G G C C C C T C C C A G C A T G C C A C T G A C C T G G

CCTCTCCCTGACACAGCCAGAACATGCCGAGCCGGAGGTCCAGGTGGTGC
CGGGGTCTGGCCAGATCATCTTCCTGCCCTTCACCTGCATTGGGTACACG
GCCACCAATCAGGACTTCATCCAGCGCCTGAGCACACTGATCCGGCAGGC
CATCGAGCGGCAGCTGCCTGCCCTGGATCGAGGCTGCCAACAGCGGGAGG
AGGGCCAGGGTGAACAGGGCGAGGAGGAGGATGAGGAGGAGGAAGAAGAG
GAGGACGTGGCTGAGAACCGCTACTTTGAAATGGGGCCCCCAGACGTGGA
GGAGGAGGAGGGAGGAGGCCAGGGGGAGGAAGAGGAGGAGGAAGAGGAGG
ATGAAGAGGCCGAGGAGGAGCGCCTGGCTCTGGAATGGGGCCTGGGGCGC
GACGAGGACTTCCTGCTGGAGCACATCCGCATCCTCAAGGTGCTGTGGTG
CTTCCTGATCCATGTGCAGGGCAGTATCCGCCAGTTCGCCGCCCTGCCTTG
TGCTCACCGACTTCGGCATCGCAGTCTTCGAGATCCCGCACCAGGAGTCT
CGGGGCAGCAGCCAGCACATCCTCTCCTCCCTGCGCTTTGTCTTTTGCTT
CCCGCATGGCGACCTCACCGAGTTTGGCTTCCTCATGCCGAGCTGTGTCT
TGGTGCTCAAGGTACGGCACAGTGAGAACACGCTCTTCATTATCTCGGAC
GCCGCCAACCTGCACGAGTTCACGCGGACCTGCGCTCATGCTTTGCACC
CCAGCACATGGCCATGCTGTGTAGCCCCATCCTCTACGGCAGCCACACCA
GCCTGCAGGAGTTCTCTGCGCCAGCTGCTCACCTTCTACAAGGTGGCTGGC
GGCTGCCAGGAGCGCAGCCAGGGCTGCTTCCCGTCTACCTGGTCTACAG
TGACAAGCGCATGGTGCAGACGCGCCCGGGGACTACTCAGGCAACATCG
AGTGGCCAGCTGCACACTCTGTTCAGCCGTGCGGCGCTCCTGCTGCGCG
CCCTCTGAGGCCGTCAAGTCCGCCGCCATCCCTACTGGCTGTTGCTCAC
GCCCCAGCCTCAACGTCAACAAGCCGACTTCAACCCCATGCCAAC
GTGGCACCACAACCTGTGCAACCGCAACAGCTTCAAGCTCAGCCGTGTG
CCGCTCTCCACCGTGCTGCTGGACCCACACGCAGCTGTACCCAGCCTCG
GGGCGCCTTTGCTGATGGCCACGTGCTAGAGCTGCTCGTGGGGTACCGCT
TTGTCACTGCCATCTTCGTGCTGCCCCACGAGAAGTTCCACTTCCTGCGC

GTCTACAACCAAGCTGCGGGCTCGCTGCAGGACCTGAAGACTGTGGTTCAT
CGCCAAGACCCCCGGGACGGGAGGCAGCCCCCAGGGCTCCTTTGCGGATG
GCCAGCCTGCCGAGCGCAGGGCCAGGTGAGATCAAGCACAGCTCTCAGGG
GCCCCGGGGGCACGGGTCTGGCATGTGTGTGATCTCAGCATCTGCGGGCTA
GTGTGGGCTGGGAGTTGCTGCGAGAGCTGGGCCCCCTCCCCCTGCCCCCT
CGCCCCCCCCGGGCCTCCCTCTACATCACCACCCCAGGTTTGGTGCCAGG
CTGCTCCTTATCTCAGTGCTGTAGAAGAAGCCCAGGAAAGCTGTCTCTCTC
ACAAATGGGTTGGCCCAGCCTCTTGCCACCCATGAAGGGCAGGCCAAGG
GGGCTGCCCCACCTTTGCCTGCCAGTGGGAGAGCAACAGGTGCAGCAC
ACCGAGGCCAGGAGAGCTGTCACCCTGGCTGCTGTCTCTCTGGGCCCCA
AGCATGGCCTCTGGGCACTACCTCCTCCAGGGTCACAGTCCCACGGATGG
CTCTGTGGGCCAGGATCTGCCTTAGGCTTCACCCACCTCAACATCTTGCT
GTGTTGTTTCAAGCTGGTCTCAAACCTTTGGGCTCAAACAATCCTCCGCCTC
AGCCTCCCAAAGTGTCTGGGATTACAGACATGAGCCACCGTGCCCCGCCGT
GCTGTTCTGTCTCCAATAGAGAAGCTGGTGAAGTCCCCAGTAACCCAG
AGGTGATGTGTGATGCACACAGTCTCCTCACTCTGAAGCTGCACATGCGA
TGTGAATCTTCATTTGGGGTCCGCTGTTAATATGGTGTTTTTTCGGGGGAT
ACAGCAATGACCAGCGTCCCCAGGAGGTCCCAGCAGAGGCTCTGGCCCCG
GCCCCAGTGAAGTCCCAGCTCCAGCCCCTGCAGCAGCCTCAGCCTCAGG
CCCAGCGAAGACTCCGGCCCCAGCAGAGGCCTCAACTTCAGCTTTGGTCC
CAGAGGAGAGCCAGTGAAGCTCCAGCCCCACCCCCAGCCGAGGCCCCCT
GCCCAGTACCCGAGTGAGCACCTCATCCAGGCCACCTCGGAGGAGAATCA
GATCCCTCGCACTTGCTGCCTGCCCGTCGCTCCGGCACGTCGCCAGCC
TGCGGGGCAGCGCCATCATCGAGCTCTTCCACAGCAGCATTGCTGAGGTA
GCGGCCCGGGTGTGGGTGCCAGCTATGGCACGGCCAGTCCTGAGGGCGAG
GCCAAGCTTGGCTTCAGGTGAGCCTCAGGTCCCTGGACTTCCTGATGTC

GGAGTCTCTCAGCTGAGCTGCTCACAGCTTTGAGGACCTGGGCAGTGAGGT
CCTGAGTTGCCCTCCCTGGCCATTGTGTCTGTGTCAACACCTCCTGTGCC
ACTTCCAGCCCCAGGTAGACCTCCCACCAACAGCCATCTCCCACCCCTCT
CTTCTCTCTGCTTGAAGCATACGGATTCACTGGTGAGCCAAGAGGGGC
TTCCCATGTCTCCTTGTGGAAGCTGTGGGCATGTCCCTGGTATGTGCAGG
TTGCTAGGGTGGTGGAGCTGACAGGAGGCCCCCGTCTTCAGGTTGAAAA
CGAGGAGCTGAGGCACCTCATGTGGTCTCGGTGGTGTCTACCAGACCC
CAGGGCTGGAGGTGACTGCCGTGCTGCTCTCCACCAAGGCTGTGTAC
TTTGTGCTCCACGACGGCCTCCGCCGCTACTTCTCAGAGCCAATGCAGGG
TAGGCACAGGGCCTGCTGGGGCTCAGGAGCTTGAGTGTGTGGTTGGGGC
AGGCCTGGGGGGTCACTCTCTGGAGCCAGCTGTGTGGCTTCAGGCAGCAG
TCAGCGACTTGCTGCTGAGTGGGCTGAGAGTTCCTTGTCTGAGGAAGGGAG
CTGTCTATGAGGGAGGGGTCCATGGCCAGATGTGAACGAGAATGCACTGA
GCCAGGGCCTGGTGACTGCTTGGGAACAGCCTGTGATGAGAAGGGGTTAG
GCAGCCTTTGCCCTGGGGCTGCACAGGAAGCCCTAGCCAGCGACCTGGT
GACTCCCCTGAGCTGGAAGAGGCTCAGACTCCAGAGGGCATTGCCTATGG
GGCTTTGCACGGGTGGAAGCCAGGCCAGCCAAGAGGACCTGTTCCTGCTG
GATGTGCTGCACACCTAGGAACCTTGTGCTTGCCTGCCACCGCCTCCCTC
TGTCCTTTCTCCATCACACAGATTCTGGCATCAGAAAAACACCGACTA
CAACAACAGCCCTTTCCACATCTCCCAGTGCTTCGTGCTAAAGCTTAGTG
ACCTGCAGTCAGTCAATGTGGGGCTTTTCGACCAGCATTTCCGGCTGACG
GGTGGGTGACCCCTCTGTGCTTTGTCTATTTTCGGGTGAAGGCCAGCATCA
CCAGTGGGCTTCCACCTTCCGTACGTGGGTGGGTTATCATAGACAGTTAT
CTCTGTGCTCAAGAGCCACTTCTTACCCGGGGTGGGAGGAAGCAGCTTCA
GGAAGTCTGAGAGAGCAGAACTCACGCTCCAGGGCTCAGAGCAGGAGGT
AGGGTGTGCGGCAAGCGCTGCCCCGGACAGAAGCAGAGTGGCCCTGGTC

TCGGGCAGGATGTTTCTGACTCACATTTCTGAGGAGAGAAAAGCTAAGCT
 CTTTGCCCTAATGTCTCTGTCTCCCCCTTCCAGAAAAATGCCTCAGCTCTTC
 CGGCCTGAAGGAATGGCCTCCTCCCGGGCCCATGATTCTTTCCTGTGTG
 GGCCTCCTGGCCCTGGCCTCTGGGCTGAGGCTTGCTAGGGACTCGGGGT
 GGCTCTAAGGGGCAGGGATAGGGCTGGGGAGCGCCGCCTGTGGCCCTGA
 CCAGCCCCTTCTCGTGCAGGTTCACCCCGATGCAGGTGGTACGTGCTT
 GACGCGGGACAGCTACCTGACGCACTGCTTCTCCAGCACCTCATGGTCG
 TGCTGTCTCTCTGGAACGCACGCCCTCGCCGGAGCCTGTTGACAAGGAC
 TTCTACTCCGAGTTTGGGAACAAGACCACAGGTACCCCTGTCTAGCTCAG
 GCTGCAGACAGGCTGCCTGGACAGACGTCATGGGCCCCAGGGTGGCTCTC
 GTGCCCCCAGAACCCTCTCTGCCTCTATGTCTCTCTTTTCTCACTTAGCT
 GGCCAGGGTTTTATGTGGGGCTTTTCGATGGCAGAGTCTCCACTCCAGCA
 GTCCCTCAACCATCTGGCAGACACATCTCCAGTGCCTGCTTTGGGCTCCT
 GGCTGTGGGCCCCACACTTGAGCATCCTCTCCTGCCTGTCTCATGCCG
 GGGTCTCTCGGTTGGCTTGGGGCCCTTGGTGCTCCAGCCCCACCAGGGG
 CCGGTTCCAGGCTATAGCCCAGGTGGCATCTCTCTGCAGGGAAGATGGAG
 AACTACGAGCTGATCCACTCTAGTCGCGTCAAGTTTACCTACCCCAGTGA
 GGAGGAGATTGGGGACCTGACGTTCACTGTGGCCCAAAGATGGCTGAGC
 CAGAGAAGGCCCCAGCCCTCAGCATCCTGCTGTACGTGCAGGCCTTCCAG
 GTGGGATGCCACCCCTGGGTGCTGCAGGGGCCCTGCGCCCCAAGAC
 ACTCCTGCTCACCAGCTCCGAGATCTTCTCCTCGGATGAGGACTGTGTCC
 ACTACCCACTGCCCGAGTTTGCCAAAGAGCGCCGCAGAGAGACAGGTAC
 CGGCTGGACGATGGCCGCCGCTCCGGGACCTGGACCGAGTGTCTATGGG
 CTACCAGACCTACCCGAGGCCCTCACCCCTCGTCTTCGATGACGTGCAAG
 GTCATGACCTCATGGGCAGTGTACCCCTGGACCACCTTTGGGGAGGTGCCA
 GGTGGCCCGCTAGAGCCAGCCAGGGCCGTGAAGTCCAGTGGCAGGTGTT

TGTCCCCAGTGTCTGAGAGCAGAGAGAAGCTCATCTCGCTGTTGGCTCGCC
 AGTGGGAGGCCCTGTGTGGCCGTGAGCTGCCGTGTCGAGCTCACC GGCTAG
 CCCAGGCCACAGCCAGCCTGTCTGTCTCCAGCCTGACGCCCTACTGGGGCAG
 GGCAGCAGGCTTTTGTGTTCTCTAAAAATGTTTATCCTCCCTTTGGTAC
 CTTAATTTGACTGTCTCTCGCAGAGAATGTGAACATGTGTGTGTGTGTGT
 TAATTCCTTCTCATGTTGGGAGTGAGAATGCCGGGCCCTCAGGGCTGTCT
 GGTGTGCTGTGACCTCCCACAGGTGGTACAGCCGTGCACACCAGTGTCTG
 TGTCTGTGTTGTGGGACCGTTGTTAACACGTGACACTGTGGGTCTGACT
 TTCTCTTCTACACGTCCTTTCCTGAAGTGTGAGTCCAGTCCCTTTGTTC
 TGTGTGCTGTTGCTGTTGCTGTTGCTGTTGGCATCTTGCTGCTAATCCTGA
 GGCTGGTAGCAGAATGCACATTGGAAGCTCCCAACCCATATTGTTCTTCA
 AAGTGGAGGTCTCCCTGATCCAGACAAGTGGGAGAGCCGTGGGGGCAG
 GGGACCTGGAGCTGCCAGCACCAGCGTGATTCCGTGCTGCCTGTATTCTC
 TATTCCAATAAAGCAGAGTTTGACACCGTCTGCATCTTCTAAACCAAGGG
 TCACTGGGATCGAGTCGACGGCCCTATAGTGAAGTGGTATTAGAGCTCGCG
 GCCGCGAGCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATC
 TGCAGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCCAC
 AGAGGTGGTCGATGAGGGCTGCCCTTTCCACATCCTTAGTAGGGGGTTC
 AAGATGACCCAGACTGTGCCCTGGGGAGCTTGAGCCATGCGGGAGGAT
 GAGCCATGTGCTGGAGGAGAACAGGGTAGGATGGTGTGGGGCTTTGTAG
 ACTGTCTAGAAGCAAAGAAGGTCTGCAGTGGAGGTGGTGTCTGAGGTGAA
 TCTCGAAGGTGAATAGGAGTTGAACGTTAGCAGGCAGAGGGTGGATTGCA
 GGAGAGCAGCGCCTGGGCAGGTGCCAGCGTGGCCCATCAGGGTGTCTTC
 ATGCATGGCTGTGTGCTTGCCATCCTTCCCTGCCTGCCTACCCCTGCTGC
 TTCGCTTCATGGGGCGTTTGAGCTTGGGCCACCTGCCTGCCTCGCTTG
 TGGGCAGAGGACCAAGCTGTGTGAGTTGTCCTGTCCCGGGGAGCAGCTG

AACTGGTCCGGGGTCTCGAACTGTGGGGGCTCAAAGGACTCCGGGGTCAT
TTCACGTGGGGCTGTGCCGATTCTGGGGGCTGTNGGAANGTAAAGGCCT
AAAGGGGCTCCCTGGTTANGGCCCTCAANTTTAANAACCTGGGGCCGGGG
CCCGGAATTGCCCCCAANTTTGTTTCAACNCCCCTTGGCCTTNGGCNGGG
GCAAATTTCCANGGGGAACCAATGNTTTCGCCCAAAAANGGGGCCNTTT
TAACCCNTTTCCAAANTTTGGGNCCTAAAAAAGGGTGGANTTCTCTGAANG
GG

(22) INFORMATION FOR SEQ ID NO: 22

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1070 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 22

VCLDDTVTTEKELDTVEVLKAIQKAKEVKSLSNPEKKGEDSRLSAAPCIRPSSSPPTVAPASA
SLPQPILSNQGIMFVQEEALASSLSSTDLSLTPHQPIAQCCSDSLESIPAGQAASDDLDRDVP
GAVGGASPEHAEPVQVVPQSGQIIIFLPFTCIGYTATNQDFIQLSTLIRQAIERQLPAWIEAANQRE
EGQGEQGEEDDEEEEEEDVAENRYFEMGPDPVEEEGGGQGEETEEEEDEEAEEERLALWAL
GADEDFLEHIRILKVLWCFLIHVQGSIRQFAACLVLTDFGIAVFEIPHQESRGSSQHILSSLR
FVFCFPHGDLTEFGFLMPELCLVLKVRHSENTLFIISDAANLHEFHADLRSCFAPQHMMMLCSPIL
YGSHTSLQEFRLQLLTFYKVAGGCQERSQGCFFVYLVYSDKRMVQTAAGDYSGNIEWASCTLC
SAVRRSCCAPSEAVKSAAIPLYLLTPQHLNVIKADFNMPMPNRGTHNCRNRNSFKLSRVPLSTVLLD
PTRSCTQPRGAFADGHVLELLVGYRFVTAIFVLPHEKFHFLRVYNQLRASLQDLKTIVVIAKTPGT
GGSPQGSFADGQPAERRASNDQRQEVPAEALAPAPVEVPAPAPAAASASGPAKTPAPAEASTSA
LVPETTPVEAPAPPAPAEAPQYPSHELIQATSEENQIPSHLPACPSLRHVASLRGSIIELFHSS

IAEVENEELRHLMWSSVVFYQTPGLEVTACVLLSTKAVYFVLHDGLRRYFSEPLQDFWHQKNTDY
NNSPFHISQCFVLKLSDLQSVNVGLFDQHFRLTGSTPMQVVTCLTRDSYLTHCFLQHLMVVLSSL
ERTPSPEPVDKDFYSEFGNKTGKMENYELIHSSRVKFTYPSEEEIGDLTFTVAQKMAEPEKAPA
LSILLVQAFQVGMPPPGCCRGPLRPKTLTLLTSSEIFLLDEDCVHYPLPEFAKEPPQRDRYRLDD
GRRVRDLDRVLMGYQTYPQALTLVFDDVQGHDLMGSVTLDFHGEVPGGPÀRASQGREVQWQVFP
SAESREKLISLLARQWEALCGRELPVELTG

WHAT IS CLAIMED IS:

CLAIMS

1. A DNA molecule encoding for a polypeptide including an amino acid sequence which is receptive to imidazoline compounds, said DNA molecule containing a DNA sequence with at least 75% sequence similarity with the DNA sequence shown in SEQ ID No. 4.

2. A DNA molecule according to claim 1, containing a DNA sequence with at least 75% sequence similarity with the DNA sequence shown in SEQ ID No. 2.

3. A DNA molecule according to claim 2, containing a DNA sequence with at least 75% sequence similarity with the DNA sequence of SEQ ID No. 3.

4. A DNA molecule according to claim 3, containing a DNA sequence with at least 75% sequence similarity with the DNA sequence of SEQ ID No. 1.

5. A DNA molecule according to any one of claims 1 to 4, containing a DNA sequence with at least 80% sequence similarity with the sequence of said SEQ ID No.

6. A DNA molecule according to any one of claims 1 to 4, containing a DNA sequence with at least 85% sequence similarity with the sequence of said SEQ ID No.

7. A DNA molecule according to any one of claims 1 to 4, containing a DNA sequence with at least 90% sequence similarity with the sequence of said SEQ ID No.

8. A DNA molecule according to any one of claims 1 to 4, containing a DNA sequence with at least 95% sequence similarity with the sequence of said SEQ ID No.

9. A DNA molecule according to claim 1, which is deposited with the ATCC under deposit accession no. ATCC 209217.

10. A genomic DNA molecule encoding for a polypeptide including an amino acid sequence which is receptive to imidazoline compounds, and wherein exon portions of said genomic DNA molecule include the DNA sequence as defined in claim 1.

11. A genomic DNA molecule according to claim 10, which is deposited with the ATCC under deposit accession no. ATCC 209216.

12. A 1110 bp ApaI-EcoRI restriction fragment of the DNA molecule according to claim 1.

13. A 1.85 kb EcoRI restriction fragment of the DNA molecule according to claim 4.

14. A vector containing a DNA sequence as defined in any one of claims 1-13.

15. A host cell transfected with a vector as defined in claim 14.

16. An isolated polypeptide including a site which is receptive to imidazoline compounds, said polypeptide containing an amino acid sequence with at least 80% sequence similarity with the amino acid sequence shown in SEQ ID No. 6.

17. A polypeptide as defined in claim 16, having a molecular weight of about 35 to 45 kDa.

18. A polypeptide as defined in claim 17, having a molecular weight of about 37 kDa.

19. An isolated polypeptide including a site which is receptive to imidazoline compounds, said polypeptide containing an amino acid sequence with at least 80% sequence similarity with the amino acid sequence shown in SEQ ID No. 5.

20. A polypeptide as defined in claim 19, having a molecular weight of about 60 to 85 kDa.

21. A polypeptide as defined in claim 20, having a molecular weight of about 70 kDa.

22. A fragment of the amino acid sequence shown in SEQ ID No. 5 or 6, which fragment is receptive to imidazoline compounds.

23. A polypeptide according to any one of claims 16 to 22, which is immunoreactive with at least one of Reis antiserum and Dontenwill antiserum.

24. A polypeptide according to any one of claims 16 to 23, which is a human polypeptide.

25. A method of producing an isolated polypeptide including an amino acid sequence which is receptive to imidazoline compounds, said method comprising:

transfecting a host cell with a vector as defined in claim 14; and

culturing the transfected host cell in a culture medium to express the polypeptide.

26. An isolated polypeptide including an amino acid sequence which is receptive to imidazoline compounds, which polypeptide is expressed by the method of claim 25.

27. A method of screening for a ligand of an imidazoline receptor, which method comprises:

culturing a host cell as defined in claim 15 in a culture medium to express a polypeptide including an amino acid sequence which is receptive to imidazoline compounds;

contacting said polypeptide with a labelled ligand for the imidazoline receptor under conditions effective to bind the labelled ligand thereto;

contacting said polypeptide with a candidate ligand; and

detecting any displacement of the labelled ligand from said polypeptide, wherein displacement signifies that the candidate ligand is a ligand for the imidazoline receptor.

28. The method of claim 27, wherein said contacting steps are performed in an intact cultured host cell.

29. The method of claim 27, further comprising isolating the cell membrane of said cultured host cell prior to performing said contacting steps.

30. The method of claim 27, wherein said contacting of said imidazoline receptive polypeptide with said candidate ligand is conducted at a plurality of candidate ligand concentrations.

31. The method of claim 27, wherein the labelled ligand is radiolabelled.

32. A method of obtaining a DNA material encoding a polypeptide which is receptive to imidazoline compounds, said method comprising:

providing a labelled DNA probe by labelling a DNA molecule identical or complementary to a DNA molecule as defined in any one of claims 1 to 9 or a restriction fragment thereof;

contacting said DNA probe with genetic material suspected of encoding said imidazoline receptive polypeptide;

hybridizing said DNA probe and said genetic material under stringent hybridization conditions;

identifying any portion of the genetic material which hybridizes to said DNA probe; and

isolating said identified material.

33. A method according to claim 32, wherein the genetic material is derived from a library selected from the group consisting of RNA library, cDNA library and genomic DNA library.

34. A method according to claim 33, wherein said library is a human library.

35. A method according to claim 32, wherein the labelled DNA probe is provided by labelling a restriction fragment according to claim 12 or 13.

36. A method of raising antibodies immunoreactive with a polypeptide which is receptive to an imidazoline compound, which method comprises:

injecting an animal with a polypeptide as defined in any one of claims 16 to 24 and 26; and

isolating antibodies produced by the animal.

ABSTRACT

A genomic DNA encoding a human imidazoline receptor is described. cDNAs encoding the receptor and fragments thereof are also provided. An amino acid sequence predicted to be 120,000 MW for nearly the entire protein is identified, as well as a middle fragment believed to contain the imidazoline binding site of the receptor. The protein is highly unique in its sequence and may represent the first in a novel family of receptor proteins.

Methods of cloning the cDNA and expressing the imidazoline receptor in a host cell are described. Methods of preparing antibodies against the transfected protein are also described. Also, a screening method for identifying additional subtypes of this receptor are identified. Also, screening methods for identifying drugs that interact with the imidazoline receptor are described.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: JOHN E. PILETZ ET AL.

Serial No.: Divisional of
08/922,635

Prior Group Art Unit: 1646

Filed: OCTOBER 8, 1999

Prior Examiner: E. Lazar-Wesley

Title: DNA MOLECULES ENCODING IMIDAZOLINE RECEPTIVE
POLYPEPTIDES AND POLYPEPTIDES ENCODED THEREBY

PROPOSED SUBSTITUTE DRAWINGS

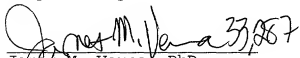
Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Applicants hereby propose to amend the application drawings
by deleting the Figure 6 originally filed with the application,
and substituting therefor the attached new Figure 6A-6B.

Respectfully submitted,

October 8, 1999

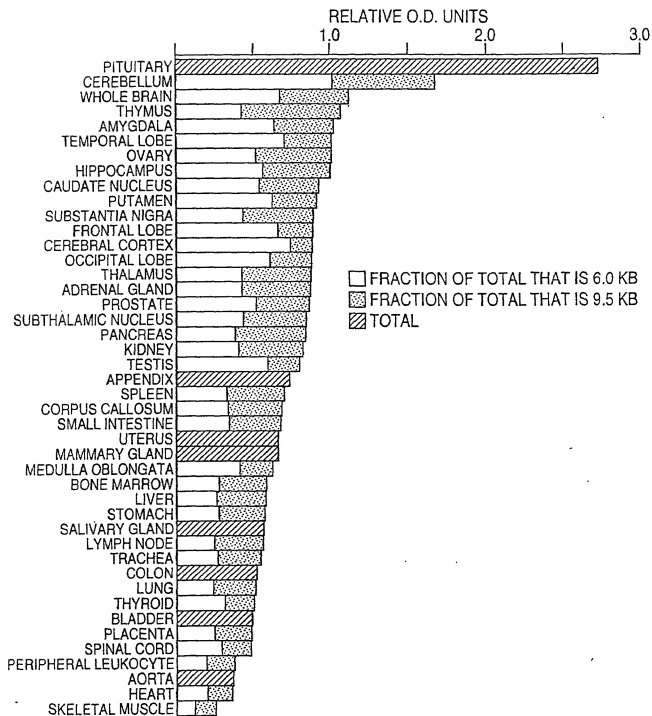

James M. Verna, PhD
Registration No. 33,287

J. D. Evans
Registration No. 26,269

EVENSON, McKEOWN, EDWARDS
& LENAHAAN, P.L.L.C.
1200 G Street, N.W., Suite 700
Washington, DC 20005
Telephone No.: (202) 628-8800
Facsimile No.: (202) 628-8844

JDE:JMV:atc

FIG. 6A



0014543.100899

FIG. 6B

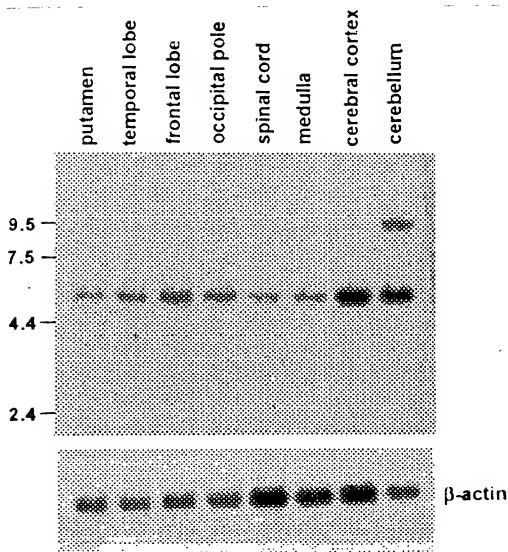


FIG. 1A

HIPPOCAMPUS

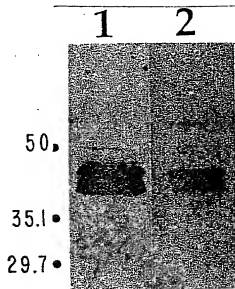


FIG. 1B

NRL

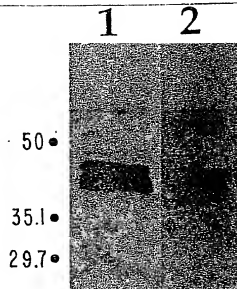
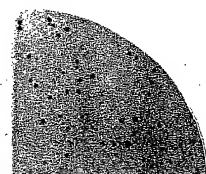
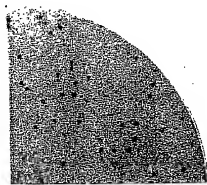


FIG. 2A



REIS AB
1:15,000 DILUTION

FIG. 2B



DONTENWILL AB
1:20,000 DILUTION

FIG. 3

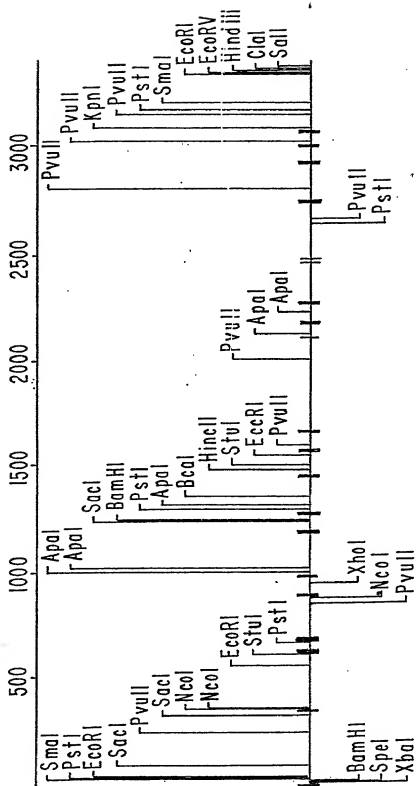


FIG. 4

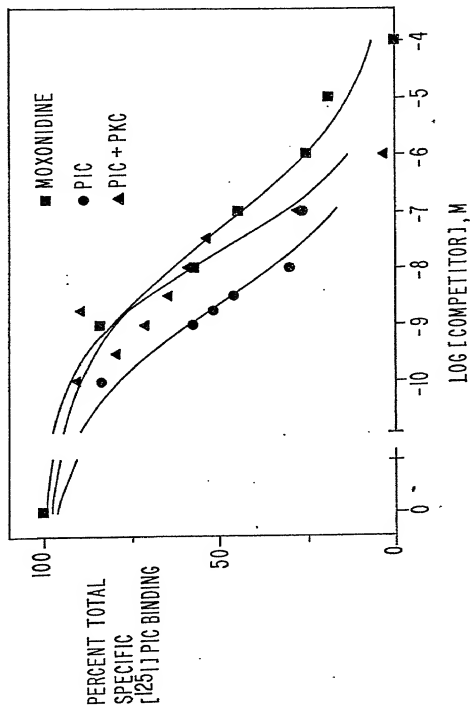
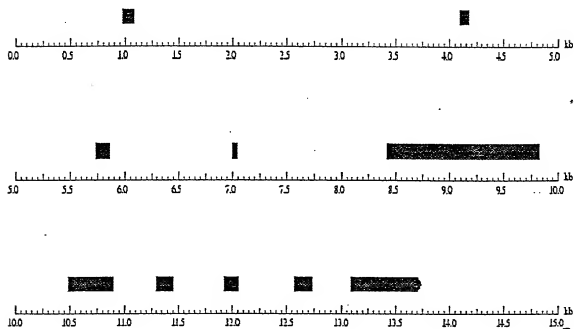


FIG. 5

GENSCAN predicted genes in sequence gs_seqfile

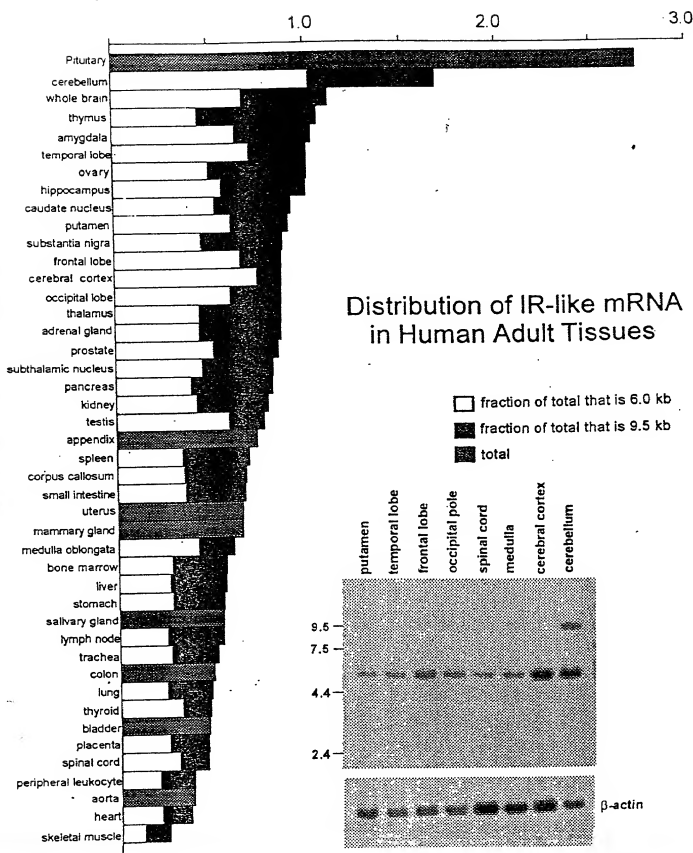


15.0 kb

Key:  Initial exon  Internal exon  Terminal exon  Single-exon gene  Optimal exon  Suboptimal exon

FIG. 6

Relative O.D. Units



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: JOHN E. PILETZ ET AL.

Serial No.: 08/922,635 Group Art Unit: 1646

Filed: SEPTEMBER 3, 1997 Examiner: E. Lazar-Wesley

Title: DNA MOLECULES ENCODING IMIDAZOLINE RECEPTIVE
POLYPEPTIDES AND POLYPEPTIDES ENCODED THEREBY

SUBSTITUTE POWER OF ATTORNEY

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

The assignee, The University of Mississippi Medical Center, was assigned the right to execute all papers in connection with the above-captioned application and any continuing applications thereof. This assignment was recorded at Reel 8839, Frames 753-765, on December 8, 1997.


The undersigned authorized agent of the assignee of the above-captioned application hereby revokes all prior powers of attorney and appoints as its attorneys Herbert I. Cantor, Reg. No. 24,392; James F. McKeown, Reg. No. 25,406; Donald D. Evenson, Reg. No. 26,160; Joseph D. Evans, Reg. No. 26,269; Gary R. Edwards, Reg. No. 31,824; Jeffrey D. Sanok, Reg. No. 32,169; and James M. Verna, Reg. No. 33,287; to prosecute the above-captioned application and transact all business in the United States Patent and Trademark Office in connection therein with.

Please address all future correspondence relating to the above-captioned application to:

J. D. Evans
EVENSON, McKEOWN, EDWARDS & LENAHA, P.L.L.C.
1200 G Street, N.W., Suite 700
Washington, D.C. 20005-3814

Address telephone communications to J.D. Evans at (202) 628-8800.

Date: 7/7/99


A. Wallace Conerly, M.D.
Vice Chancellor for Health Affairs for
The University of Mississippi Medical Center



DECLARATION AND POWER OF ATTORNEY FOR U.S. PATENT APPLICATION

(X) Original () Supplemental () Substitute () PCT () DESIGN

As a below named inventor, I hereby declare that: my residence, post office address and citizenship are as stated below next to my name; that I verily believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural inventors are named below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

Title: DNA MOLECULES ENCODING IMIDAZOLINE RECEPTIVE POLYPEPTIDES AND POLYPEPTIDES ENCODED THEREBY

of which is described and claimed in:

- () the attached specification, or
 (X) the specification in the application Serial No. _____ filed September 3, 1997;
 and with amendments through _____ (if applicable), or
 () the specification in International Application No. PCT/_____, filed _____, and as amended on _____ (if applicable).

I hereby state that I have reviewed and understand the content of the above-identified specification, including the claims, as amended by any amendment(s) referred to above.

I acknowledge my duty to disclose information of which I am aware which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, §119 (and §172 if this application is for a Design) of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

COUNTRY	APPLICATION NO.	DATE OF FILING	PRIORITY CLAIMED

I hereby claim the benefit under Title 35, United States Code §119(c) and §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

APPLICATION SERIAL NO.	U.S. FILING DATE	STATUS: PATENTED, PENDING, ABANDONED
08/650,766	May 20, 1996	Pending
60/012,600	March 1, 1996	Pending

And I hereby appoint V. M. Creedon, Reg. No. 17,111; John T. Miller, Reg. No. 21,120; Michael R. Davis, Reg. No. 25,134; Matthew M. Jacob, Reg. No. 25,154; Jeffrey Nolton, Reg. No. 25,408; Warren M. Check, Jr., Reg. No. 33,367; Nils E. Pedersen, Reg. No. 33,145; and Charles R. Watts, Reg. No. 33,142, who together constitute the firm of WENDEROTH, LIND & PONACK, jointly and severally, attorneys to prosecute this application and to transact all business in the U.S. Patent and Trademark Office connected therewith.

I hereby authorize the U.S. attorneys named herein to accept and follow instructions from UNIVERSITY OF MISSISSIPPI MEDICAL CENTER as to any action to be taken in the U.S. Patent and Trademark Office regarding this application without direct communication between the U.S. attorneys and myself. In the event of a change in the persons from whom instructions may be taken, the U.S. attorneys named herein will be so notified by me.

Send Correspondence to

Direct Telephone Calls to:

WENDEROTH, LIND & PONACK
Southern Building - Suite 700
805 Fifteenth Street, N.W.
Washington, D.C. 20005

WENDEROTH, LIND & PONACK
Area Code (202) 371-8850

Full Name of First Inventor	FAMILY NAME PILETZ	FIRST GIVEN NAME John	SECOND GIVEN NAME E.
Residence & Citizenship	CITY Madison	STATE OR COUNTRY Mississippi	COUNTRY OF CITIZENSHIP U.S.A.
Post Office Address	ADDRESS 14 Peachtree Lane, Madison, Mississippi 39110	CITY	STATE OF COUNTRY ZIP CODE
Full Name of Second Inventor	FAMILY NAME IVANOV	FIRST GIVEN NAME Tina	SECOND GIVEN NAME R.
Residence & Citizenship	CITY Manchester	STATE OR COUNTRY England	COUNTRY OF CITIZENSHIP Canada
Post Office Address	ADDRESS c/o University of Manchester, Division of Neurosciences, School of Biological Sciences, Stopford Building, 1.124, Oxford Road, Manchester, England M13 9PT	CITY	STATE OF COUNTRY ZIP CODE
Full Name of Third Inventor	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME
Residence & Citizenship	CITY	STATE OR COUNTRY	COUNTRY OF CITIZENSHIP
Post Office Address	ADDRESS	CITY	STATE OF COUNTRY ZIP CODE
Full Name of Fourth Inventor	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME
Residence & Citizenship	CITY	STATE OR COUNTRY	COUNTRY OF CITIZENSHIP
Post Office Address	ADDRESS	CITY	STATE OF COUNTRY ZIP CODE

Full Name of Fifth Inventor	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME
Residence & Citizenship	CITY	STATE OR COUNTRY	COUNTRY OF CITIZENSHIP
Post Office Address	ADDRESS	CITY	STATE OF COUNTRY ZIP CODE

Full Name of Sixth Inventor	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME
Residence & Citizenship	CITY	STATE OR COUNTRY	COUNTRY OF CITIZENSHIP
Post Office Address	ADDRESS	CITY	STATE OF COUNTRY ZIP CODE

I further declare that all statements made herein of my own knowledge are true, and that all statements on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

1st Inventor John E. Piletz Date 9/28/97
John E. PILETZ
2nd Inventor Tina R. Ivanov Date October 7, 1997
Tina R. IVANOV
3rd Inventor _____ Date _____
4th Inventor _____ Date _____
5th Inventor _____ Date _____
6th Inventor _____ Date _____

The above application may be more particularly identified as follows:

U.S. Application Serial No. _____ Filing Date September 3, 1997
Applicant Reference Number _____ Atty Docket No. 1342/PILETZ
Title of Invention DNA MOLECULES ENCODING IMIDAZOLINE RECEPTIVE POLYPEPTIDES AND POLYPEPTIDES ENCODED THEREBY

SEQUENCE LISTING

<110> FILETZ, John E.
IVANOV, Tina R.

<120> DNA MOLECULES ENCODING IMIDALINE RECEPTIVE POLYPEPTIDES
AND POLYPEPTIDES ENCODED THEREBY

<130> Corrected Sequence Listing

<140> 08/922,635

<141> 1997-09-03

<150> 08/650,766

<151> 1996-05-20

<150> 60/012,600

<151> 1996-03-01

<160> 22

<170> PatentIn Ver. 2.0

<210> 1

<211> 3385

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1398)..(3383)

<400> 1

```
gctctagaac tagtggatcc cccgggctgc aggaattcca gtttaataact aaccctaattg 60
tgtgactgcg gtttacaaag agctctgtat cacctgggat agctttcagt agcaattcac 120
tacaactggt cctaaaaaat aataacaata ataataataa ttagagaatt aaaacccaac 180
agcatgttga atgggttaaaa tcacgtaaga actgaaattt ggggtggggg tgtcctcaac 240
agctgagctt gtcctagcag tgaaaaatgct cgcctccaag cagggtctcag aaaggtcttg 300
agccctccag gcagagggct gagctcaggg ggctcttgga ggacactcac cccatgggtcc 360
```


gcc agc agc ctc tgc tcc act gac agt ctg act ccc gag cac cag ccc	1766
Ala Ser Ser Leu Ser Ser Thr Asp Ser Leu Thr Pro Glu His Gln Pro	
110 115 120	
att gcc cag gga tgt tct gat tcc ttg gag tcc atc cct gcg gga cag	1814
Ile Ala Gln Gly Cys Ser Asp Ser Leu Glu Ser Ile Pro Ala Gly Gln	
125 130 135	
gca gct tcc gat gat tta agg gac gtg cca gga gct gtt ggt ggt gca	1862
Ala Ala Ser Asp Asp Leu Arg Asp Val Pro Gly Ala Val Gly Gly Ala	
140 145 150 155	
agc cca gaa cat gcc gag ccg gag gtc cag gtg gtg ccg ggg tct ggc	1910
Ser Pro Glu His Ala Glu Pro Glu Val Gln Val Val Pro Gly Ser Gly	
160 165 170	
cag atc atc ttc ctg ccc ttc acc tgc att ggc tac acg gcc acc aat	1958
Gln Ile Ile Phe Leu Pro Phe Thr Cys Ile Gly Tyr Thr Ala Thr Asn	
175 180 185	
cag gac ttc atc cag ccg ctg agc aca ctg atc cgg cag gcc atc gag	2006
Gln Asp Phe Ile Gln Arg Leu Ser Thr Leu Ile Arg Gln Ala Ile Glu	
190 195 200	
cgg cag ctg cct gcc tgg atc gag gct gcc aac cag cgg gag gag ggc	2054
Arg Gln Leu Pro Ala Trp Ile Glu Ala Ala Asn Gln Arg Glu Glu Gly	
205 210 215	
cag ggt gaa cag ggc gag gag gag gat gag gag gag gaa gaa gag gag	2102
Gln Gly Glu Gln Gly Glu Glu Glu Asp Glu Glu Glu Glu Glu Glu	
220 225 230 235	
gac gtg gct gag aac ccg tac ttt gaa atg ggg ccc cca gac gtg gag	2150
Asp Val Ala Glu Asn Arg Tyr Phe Glu Met Gly Pro Pro Asp Val Glu	
240 245 250	
gag gag gag gga gga ggc cag ggg gag gaa gag gag gag gaa gag gag	2198
Glu Glu Glu Gly Gly Gly Gln Gly Glu Glu Glu Glu Glu Glu Glu	
255 260 265	
gat gaa gag gcc gag gag gag ccg ctg gct ctg gaa tgg gcc ctg ggc	2246
Asp Glu Glu Ala Glu Glu Glu Arg Leu Ala Leu Glu Trp Ala Leu Gly	
270 275 280	
ccg gac gag gac ttc ctg ctg gag cac atc ccg atc ctc aag gtg ctg	2294
Ala Asp Glu Asp Phe Leu Leu Glu His Ile Arg Ile Leu Lys Val Leu	
285 290 295	
tgg tgc ttc ctg atc cat gtg cag ggc agt atc ccg cag ttc gcc gcc	2342
Trp Cys Phe Leu Ile His Val Gln Gly Ser Ile Arg Gln Phe Ala Ala	
300 305 310 315	
tgc ctt gtg ctc acc gac ttc ggc atc gca gtc ttc gag atc ccg cac	2390
Cys Leu Val Leu Thr Asp Phe Gly Ile Ala Val Phe Glu Ile Pro His	
320 325 330	
cag gag tct ccg ggc agc agc cag cac atc ctc tcc tcc ctg ccg ttt	2438
Gln Glu Ser Arg Gly Ser Ser Gln His Ile Leu Ser Ser Leu Arg Phe	
335 340 345	

gtc ttt tgc ttc ccg cat ggc gac ctc acc gag ttt ggc ttc ctc atg Val Phe Cys Phe Pro His Gly Asp Leu Thr Glu Phe Gly Phe Leu Met 350 355 360	2486
ccg gag ctg tgt ctg gtg ctc aag gta cgg cac agt gag aac acg ctc Pro Glu Leu Cys Leu Val Leu Lys Val Arg His Ser Glu Asn Thr Leu 365 370 375	2534
ttc att atc tcg gac gcc gcc aac ctg cac gag ttc cac gcg gac ctg Phe Ile Ile Ser Asp Ala Ala Asn Leu His Glu Phe His Ala Asp Leu 380 385 390 395	2582
cgc tca tgc ttt gca ccc cag cac atg gcc atg ctg tgt agc ccc atc Arg Ser Cys Phe Ala Pro Gln His Met Ala Met Leu Cys Ser Pro Ile 400 405 410	2630
ctc tac ggc agc cac acc agc ctg cag gag ttc ctg cgc cag ctg ctc Leu Tyr Gly Ser His Thr Ser Leu Gln Glu Phe Leu Arg Gln Leu Leu 415 420 425	2678
acc ttc tac aag gtg gct ggc gcc tgc cag gag cgc agc cag gcc tgc Thr Phe Tyr Lys Val Ala Gly Gly Cys Gln Glu Arg Ser Gln Gly Cys 430 435 440	2726
ttc ccc gtc tac ctg gtc tac agt gac aag cgc atg gtg cag acg gcc Phe Pro Val Tyr Leu Val Tyr Ser Asp Lys Arg Met Val Gln Thr Ala 445 450 455	2774
gcc ggg gac tac tca ggc aac atc gag tgg gcc agc tgc aca ctc tgt Ala Gly Asp Tyr Ser Gly Asn Ile Glu Trp Ala Ser Cys Thr Leu Cys 460 465 470 475	2822
tca gcc gtg cgg cgc tcc tgc tgc cgc ccc tct gag gcc gtc aag tcc Ser Ala Val Arg Arg Ser Cys Cys Ala Pro Ser Glu Ala Val Lys Ser 480 485 490	2870
gcc gcc atc ccc tac tgg ctg ttg ctc acg ccc cag cac ctc aac gtc Ala Ala Ile Pro Tyr Trp Leu Leu Leu Thr Pro Gln His Leu Asn Val 495 500 505	2918
atc aag gcc gac ttc aac ccc atg ccc aac cgt ggc acc cac aac tgt Ile Lys Ala Asp Phe Asn Pro Met Pro Asn Arg Gly Thr His Asn Cys 510 515 520	2966
cgc aac cgc aac agc ttc aag ctc agc cgt gtg ccg ctc tcc acc gtg Arg Asn Arg Asn Ser Phe Lys Leu Ser Arg Val Pro Leu Ser Thr Val 525 530 535	3014
ctg ctg gac ccc aca cgc agc tgt acc cag cct cgg gcc gcc ttt gct Leu Leu Asp Pro Thr Arg Ser Cys Thr Gln Pro Arg Gly Ala Phe Ala 540 545 550 555	3062
gat ggc cac gtg cta gag ctg ctc gtg ggg tac cgc ttt gtc act gcc Asp Gly His Val Leu Glu Leu Leu Val Gly Tyr Arg Phe Val Thr Ala 560 565 570	3110
atc ttc gtg ctg ccc cac gag aag ttc cac ttc ctg cgc gtc tac aac Ile Phe Val Leu Pro His Glu Lys Phe His Phe Leu Arg Val Tyr Asn 575 580 585	3158

cag	ctg	cgg	gcc	tcg	ctg	cag	gac	ctg	aag	act	gtg	gtc	atc	gcc	aag	3206
Gln	Leu	Arg	Ala	Ser	Leu	Gln	Asp	Leu	Lys	Thr	Val	Val	Ile	Ala	Lys	
		590					595					600				
acc	ccc	ggg	acg	gga	ggc	agc	ccc	cag	ggc	tcc	ttt	gcg	gat	ggc	cag	3254
Thr	Pro	Gly	Thr	Gly	Gly	Ser	Pro	Gln	Gly	Ser		Phe	Ala	Asp	Gly	Gln
		605				610					615					
cct	gcc	gag	cgc	agg	gcc	agc	aat	gac	cag	cgt	ccc	cag	gag	gtc	cca	3302
Pro	Ala	Glu	Arg	Arg	Ala	Ser	Asn	Asp	Gln	Arg	Pro	Gln	Glu	Val	Pro	
		620			625					630					635	
gca	gag	gct	ctg	gcc	ccg	gcc	cca	gtg	gaa	gtc	cca	gct	cca	gcc	ccg	3350
Ala	Glu	Ala	Leu	Ala	Pro	Ala	Pro	Val	Glu	Val	Pro	Ala	Pro	Ala	Pro	
			640					645						650		
gaa	ttc	gat	atc	aag	ctt	atc	gat	acc	gtc	gac	ct					3385
Glu	Phe	Asp	Ile	Lys	Leu	Ile	Asp	Thr	Val	Asp						
		655					660									

<210> 2
 <211> 1954
 <212> DNA
 <213> Homo sapiens

<400> 2

atgacaggcc aggtgggagc gcagacggtc tcgggtggca aaagaagcat tgcaggctctg 60
 acactttgtga ggccgctcag aagtgtgcac ctgcttgata tgtccgttca agtgatcagg 120
 cctggagaag cctttccac agctctggca gatgtaaggc ggaattcccc agagaagaag 180
 ggtggtgaag actcccggt ctcagctgcc cctgcacat gaccacagcag ctccccctccc 240
 actgtggctc ccgcatctgc ctccctgccc cagcccatcc tctctaacca aggaatcatg 300
 ttctgtcagg aggaggccct ggccagcagc ctctcgtcca ctgacagtct gactcccag 360
 caccagccca ttgcccaggg atgttctgat tccttgaggt ccattccctgc gggacaggca 420
 gcttccgatg atttaaggga cgtgccagga gctgttggtg gtgcaagccc agaacaatgcc 480
 gagccggagg tccaggtggt gccggggtct ggccagatca tcttctgtcc ctccactctgc 540
 attggctaca cggccaccaa tcaggacttc atccagcgcc tgagcacact gatccggcag 600
 gccatcgagc ggcatctgcc tgccctggatc gaggctgccca accagcggga ggagggccag 660
 ggtgaacagg gcgaggagga ggatgaggag gaggaagaag aggaggacgt ggctgagaac 720
 cgctactttg aaatggggcc ccagacagtg gaggaggagg agggaggagg ccagggggag 780
 gaagaggagg aggaagagga ggatgaagag gccgaggagg agcgccctggc tctggaatgg 840
 gccctgggag cggacgagga cttctgtgtg gaccacatcc gcattcctcaa ggtgctgtgg 900

tgcttcctga tccatgtgca gggcagttac cgcagttcgc cgcctgcct tgtgctcacc 960
 gacttcggca tgcagttctt cgagatcccg caccaggagt ctccggggcag cagccagcac 1020
 atctctctct ccttcgcgtt tgtcttttgc ttcccgcatg gcgacctcac cgagttttgc 1080
 ttctcatcgc cggagctgtg tctggtgctc aaggtacggc acagtgcgaa cactctcttc 1140
 attatctcgg acgcgcgcaa cctgcacgag ttccacgcgg acctgcgctc atgcttttga 1200
 cccagcacaa tggccatgct gtgtagcccc atctctacg gcacccacac cagcctgcag 1260
 gagtctctgc gccagctgct cactctctac aaggtggctg gcgctgcca ggagcgcagc 1320
 cagggtgct tcccgtcta cctggtctac agtgacaagc gcctggtgca gacggcggcc 1380
 ggggactact caggcaacat cgagtgggcc agctgcacac tctgttcagc cgtgcggcgc 1440
 tctgtctgcg cgcctctga ggcgctcaag tccgcgcgca tccctactg gctgttgctc 1500
 acgccccagc acctcaacgt catcaaggcc gactcaacc ccctgcgcaa ccgtggcacc 1560
 cacaactgtc gcaacgcgaa cagcttcaag ctacgcccgtg tgccgctctc caccgtgtgt 1620
 ctggacccca cagcagctgt taoccagcct cggggcgctt ttgctgatgg ccacgtgcta 1680
 gagctgctcg tggggtaccg ctttgtcact gccatctctg tgcctgccc aagaagttc 1740
 cacttctctc gcgtctacaa ccagctgcgg gcctcgtgc aggacctgaa gactgtggtc 1800
 atcgccaaga cccccgggag gggaggcagc cccagggct cttttgcgga tggccagcct 1860
 gccgagcgca gggccagcaa tgaccagcgt cccagggagg tccacgcaga ggctctggcc 1920
 ccggccccag tggaagtccc agctccagcc ccgg 1954

<210> 3
 <211> 3318
 <212> DNA
 <213> Homo sapiens

<400> 3

aattocagtt taatactaac cctaattgtg gactgcgggt tacaaagagc tctgtatcac 60
 ctgggatagc ttctcagtag aattcactac aactggctct aaaaaataat aacaataata 120
 ataataatta gagaattaaa acccaacagc atgttgaatg gttaaaatca cgtaagaact 180
 gaaatttggg gtgggggtgt cctcaacagc tgagcttgct ctacagtgga aaatgctcgc 240
 ctcaagcagc ggctcagaaa ggtctggagc cctccaggca gagggtgag ctacgggggc 300
 tcttgaggga cactcaccoc atggtccatg ggaatgctct ggcttctta aaaacagttg 360
 ggcatccgca ttgtataagt aggtggagac cctagtgtgg ttcttttgaa ggatatggga 420
 agggaggatg acgaactaga gaagtgggag gggacacaaa tcaactgaggt ccagagaata 480

catagatttg ggtataggat tggggctact aagaattgag caccaggaat tccagcttct 540
tcccattaaa gaaactggga ctgggtttgc ctggaggcc tatgtagtgt tttctgcccc 600
tgtcccatc caagtctcat tgatatctct gcagaatc agatgaaaat ctattcttaa 660
agaccatttg gagaatgggt ggtggagaag gagttggagt ggggttggg ggcagttaaa 720
aatgaataaa aatctctcag ctacagaacc caaacatcac ttcctctcgc attcacagca 780
tttccacga gtcaccagat ggttgtttcc gtggggacac agcagctgcc tcatctccct 840
tcaggcccca tgggctgctg gtcacactca ggatctacta aagatgacgc aaatgccgac 900
tgaacaatct gaaaccocaa ggaactcgagg agagacatgt tctgtgagg agagaaaggt 960
gagccaaggg cagggccagc gtcccccagg gggcccccga gagccggac atgcaccttc 1020
tggatgtgtt gtttcaagta ggacttagag cggaagaagc tcccacattc agggcatggg 1080
tactcttct ccccatcaga ctccattttg ttttgggga ctgccatgtc gcaggagaaa 1140
gagccatttg cactctgctt ctctggcgctc ttcaggtcgc tggcatctga gaggtcacca 1200
taggagtcag agctctcaat cggatcctga tgtgagcatt tctggccttc tcggttacag 1260
atactgcaga agttgtctgg cccctcgctg tgccttctca ggtggtctgc catgtatgct 1320
gcccgcgaag acttcccaca cacctggcag ggcacctgtt ctctcatgaca ggccagggtg 1380
gagcgcagac ggtctcgggt ggcaaaagaa gcattgcagg tctgacactt gtgaggccgc 1440
tcagaagtgt gcacctgctt gatattgtcc ttcaagtgat caggcctgga gaagccttcc 1500
ccacagctct ggcagatgta aggcggaatt cccagagaa gaagggtggt gaagactccc 1560
ggctctcagc tgccccctgc atcagaccca gcagctcccc tcccactgtg gctcccgcac 1620
ctgcctccct gcccccagcc atcctctcta accaaggaa catgttctgt caggaggagg 1680
ccctggccag cagcctctcg tccactgaca gtctgaactc cgagcaccag cccattgccc 1740
agggatgttc tgattccttg gagtccatcc ctgcgggaca ggcagcttcc gatgatttaa 1800
gggacgtgcc aggagctgtt ggtggtgcaa gccagaaca tgccagagccg gaggtccagg 1860
tggtgccccg gtcctggccag atcatcttcc tgcccttcac ctgcattggc tacacggcca 1920
ccaatcagga ctctcatccag cgcctgagca cactgatccg gcaggccatc gagcgccagc 1980
tgctgctctg gatcagggtc gccaaaccagc gggaggaggg ccagggtgaa caggggcagg 2040
aggagatgta ggaggaggaa gaagaggagg acgtggctga gaaccgctac ttgaaatgg 2100
ggcccccaga cgtggaggag gaggaggagg gaggccaggg ggaggaagag gaggaggaa 2160
aggagatgta agaggccgag gaggagcgcc tggctctgga atgggcccgt ggcgcggacg 2220
aggacttct gctggagcac atccgcatcc tcaagtgctg gtggtgcttc ctgatccatg 2280
tgcagggcag tatccgccag ttcccgccct gccctgtgct caccgacttc ggcacgcgac 2340

tcttcgagat cccgcaccag gagtctcggg gcagcagcca gcacatcttc tctctccctgc 2400
 gctttgtctt ttgcttcccg catggcgacc tcaccgagtt tggcttcttc atgccggagc 2460
 tgtgtctggt gctcaaggta cggcacagtgc agaacacgct cttcattatc tcggacgccc 2520
 ccaacctgca cgagttccac gcggacctgc gctcatgctt tgcaccccag cacatggcca 2580
 tgctgtgtag ccccatcttc tacggcagcc acaccagcct gcaggagtgc ctgcgccagc 2640
 tgctcacctt ctacaaggta gctggcggct gccaggagcg cagccagggc tgcctccccc 2700
 ttacacctgt ctacagtgc aagcgcatgg tgcagacggc cgccggggac taactcaggea 2760
 acatcgagtg ggccagctgc acactctgtt cagccgtgcg gcgctcctgc tgcgcgccct 2820
 ctgaggccgt caagtccgc gccatcccct actggtctgt gctcacgcc cagcacctca 2880
 acgtcatcaa ggccgacttc aaccccacgc ccaaccgtgg caccacaaac tgcgcgaacc 2940
 gcaacagctt caagctcagc cgtgtgcgcg tctccaccgt gctgctggac cccacacgca 3000
 gctgtacca gctcggggc gcctttgctg atggccacct gctagagctg ctctgggggt 3060
 accgctttgt cactgccatc ttcgtgctgc ccaacgagaa gttccaattc ctgcgcgtct 3120
 acaaccagct gcgggcctcg ctgcaggacc tgaagactgt ggtcatcgcc aagacccccg 3180
 ggacgggagg cagccccag ggctcctttg cggatggcca gctgcgcgag cgcagggccca 3240
 gcaatgacca gcgtcccccag gaggtcccag cagaggctct ggcgccggcc ccagtggaag 3300
 tccagctcc agccccgg 3318

<210> 4
 <211> 1171
 <212> DNA
 <213> Homo sapiens

<400> 4

gaggaggagg aagaggagga tgaaggaggc gaggaggagc gcctggctct ggaatggggc 60
 ctggggcgcg acgaggactt cctgctggag cacatccgca tctcaagggt gctgtggtgc 120
 ttctgatcc atgtgcaggg cagtatccgc cagttcgccg cctgcttctg gctcaccgac 180
 ttccgcatcg cagtcttcga gatcccgac caggagtctc ggggcagcag ccagcacatc 240
 ctctctctcc tgcgctttgt cttttgcttc ccgcatggcg acctcaccga gtttggtctc 300
 ctctgcgcg agctgtgtct ggtgtctcaag gtacggcaca gtgagaacac gctcttcatt 360
 atctcggagc ccgccaacct gcacgagttc caccgggacc tgcgtcatg ctttgacacc 420
 cagcacatg ccatgtgtg tagcccatc ctctacggca gccacaccag cctgcaggag 480
 ttctcgcgcc agctgtctca cttctacaag gtggctggcg gctgccagga gcgcagccag 540

ggctgcttcc cgtctacct ggctacagtc gacaagcgca tgggtgcagac ggcgcgcggg 600
 gactactcag gcaacatcga gtggggccagc tgcacactct gttcagccgt gcggcgctcc 660
 tgctgcgcgc cctctgagcg cgtcaagtcc gccgccatcc cctactggct gttgctcacg 720
 cccagcacc tcaacgtcat caaggccgac ttcaacccca tgcccaaccg tggcaccacc 780
 aactgtgcga accgcaacag ctcaagctc agccgtgtgc cgtctctccac cgtgctgctg 840
 gacccacac gcagctgtac ccagcctcgg ggccgctttg ctgatggcca cgtgctagag 900
 ctgctcgtgg ggtaccgctt tgtcactgcc atctctgtgc tgccccacga gaagttccac 960
 ttctctgcgc tctacaacca gctgcggggc tcgctgcagg acctgaagac tgtggtcacc 1020
 gccaaagacc ccgggacggg aggcagcccc cagggtcct ttgcggatgg ccagcctgcc 1080
 gagcgcagg ccagcaatga ccagcgtccc caggaggctc cagcagaggc tctgccccg 1140
 gccccagtgg aagtccacc tccagccccg g 1171

<210> 5
 <211> 651
 <212> PRT
 <213> Homo sapiens

<400> 5

Met	Thr	Gly	Gln	Val	Gly	Ala	Gln	Thr	Val	Ser	Gly	Gly	Lys	Arg	Ser
1				5					10					15	
Ile	Ala	Gly	Leu	Thr	Leu	Val	Arg	Pro	Leu	Arg	Ser	Val	His	Leu	Leu
			20					25					30		
Asp	Met	Ser	Val	Gln	Val	Ile	Arg	Pro	Gly	Glu	Ala	Phe	Pro	Thr	Ala
		35					40					45			
Leu	Ala	Asp	Val	Arg	Arg	Asn	Ser	Pro	Glu	Lys	Lys	Gly	Gly	Glu	Asp
	50					55					60				
Ser	Arg	Leu	Ser	Ala	Ala	Pro	Cys	Ile	Arg	Pro	Ser	Ser	Ser	Pro	Pro
	65				70					75				80	
Thr	Val	Ala	Pro	Ala	Ser	Ala	Ser	Leu	Pro	Gln	Pro	Ile	Leu	Ser	Asn
				85					90					95	
Gln	Gly	Ile	Met	Phe	Val	Gln	Glu	Glu	Ala	Leu	Ala	Ser	Ser	Leu	Ser
			100					105						110	
Ser	Thr	Asp	Ser	Leu	Thr	Pro	Glu	His	Gln	Pro	Ile	Ala	Gln	Gly	Cys
		115					120				125				
Ser	Asp	Ser	Leu	Glu	Ser	Ile	Pro	Ala	Gly	Gln	Ala	Ala	Ser	Asp	Asp
	130					135					140				
Leu	Arg	Asp	Val	Pro	Gly	Ala	Val	Gly	Gly	Ala	Ser	Pro	Glu	His	Ala
	145				150					155					160

Glu Pro Glu Val Gln Val Val Pro Gly Ser Gly Gln Ile Ile Phe Leu
 165 170 175
 Pro Phe Thr Cys Ile Gly Tyr Thr Ala Thr Asn Gln Asp Phe Ile Gln
 180 185 190
 Arg Leu Ser Thr Leu Ile Arg Gln Ala Ile Glu Arg Gln Leu Pro Ala
 195 200 205
 Trp Ile Glu Ala Ala Asn Gln Arg Glu Glu Gly Gln Gly Glu Gln Gly
 210 215 220
 Glu Glu Glu Asp Glu Glu Glu Glu Glu Asp Val Ala Glu Asn
 225 230 235 240
 Arg Tyr Phe Glu Met Gly Pro Pro Asp Val Glu Glu Glu Glu Gly Gly
 245 250 255
 Gly Gln Gly Glu Glu Glu Glu Glu Glu Asp Glu Glu Ala Glu
 260 265 270
 Glu Glu Arg Leu Ala Leu Glu Trp Ala Leu Gly Ala Asp Glu Asp Phe
 275 280 285
 Leu Leu Glu His Ile Arg Ile Leu Lys Val Leu Trp Cys Phe Leu Ile
 290 295 300
 His Val Gln Gly Ser Ile Arg Gln Phe Ala Ala Cys Leu Val Leu Thr
 305 310 315 320
 Asp Phe Gly Ile Ala Val Phe Glu Ile Pro His Gln Glu Ser Arg Gly
 325 330 335
 Ser Ser Gln His Ile Leu Ser Ser Leu Arg Phe Val Phe Cys Phe Pro
 340 345 350
 His Gly Asp Leu Thr Glu Phe Gly Phe Leu Met Pro Glu Leu Cys Leu
 355 360 365
 Val Leu Lys Val Arg His Ser Glu Asn Thr Leu Phe Ile Ile Ser Asp
 370 375 380
 Ala Ala Asn Leu His Glu Phe His Ala Asp Leu Arg Ser Cys Phe Ala
 385 390 395 400
 Pro Gln His Met Ala Met Leu Cys Ser Pro Ile Leu Tyr Gly Ser His
 405 410 415
 Thr Ser Leu Gln Glu Phe Leu Arg Gln Leu Leu Thr Phe Tyr Lys Val
 420 425 430
 Ala Gly Gly Cys Gln Glu Arg Ser Gln Gly Cys Phe Pro Val Tyr Leu
 435 440 445
 Val Tyr Ser Asp Lys Arg Met Val Gln Thr Ala Ala Gly Asp Tyr Ser
 450 455 460
 Gly Asn Ile Glu Trp Ala Ser Cys Thr Leu Cys Ser Ala Val Arg Arg
 465 470 475 480
 Ser Cys Cys Ala Pro Ser Glu Ala Val Lys Ser Ala Ala Ile Pro Tyr
 485 490 495

Trp Leu Leu Leu Thr Pro Gln His Leu Asn Val Ile Lys Ala Asp Phe
 500 505 510
 Asn Pro Met Pro Asn Arg Gly Thr His Asn Cys Arg Asn Arg Asn Ser
 515 520 525
 Phe Lys Leu Ser Arg Val Pro Leu Ser Thr Val Leu Leu Asp Pro Thr
 530 535 540
 Arg Ser Cys Thr Gln Pro Arg Gly Ala Phe Ala Asp Gly His Val Leu
 545 550 555 560
 Glu Leu Leu Val Gly Tyr Arg Phe Val Thr Ala Ile Phe Val Leu Pro
 565 570 575
 His Glu Lys Phe His Phe Leu Arg Val Tyr Asn Gln Leu Arg Ala Ser
 580 585 590
 Leu Gln Asp Leu Lys Thr Val Val Ile Ala Lys Thr Pro Gly Thr Gly
 595 600 605
 Gly Ser Pro Gln Gly Ser Phe Ala Asp Gly Gln Pro Ala Glu Arg Arg
 610 615 620
 Ala Ser Asn Asp Gln Arg Pro Gln Glu Val Pro Ala Glu Ala Leu Ala
 625 630 635 640
 Pro Ala Pro Val Glu Val Pro Ala Pro Ala Pro
 645 650

 <210> 6
 <211> 390
 <212> PRT
 <213> Homo sapiens

 <400> 6
 Glu Glu Glu Glu Glu Glu Asp Glu Glu Ala Glu Glu Glu Arg Leu Ala
 1 5 10 15
 Leu Glu Trp Ala Leu Gly Ala Asp Glu Asp Phe Leu Leu Glu His Ile
 20 25 30
 Arg Ile Leu Lys Val Leu Trp Cys Phe Leu Ile His Val Gln Gly Ser
 35 40 45
 Ile Arg Gln Phe Ala Ala Cys Leu Val Leu Thr Asp Phe Gly Ile Ala
 50 55 60
 Val Phe Glu Ile Pro His Gln Glu Ser Arg Gly Ser Ser Gln His Ile
 65 70 75 80
 Leu Ser Ser Leu Arg Phe Val Phe Cys Phe Pro His Gly Asp Leu Thr
 85 90 95
 Glu Phe Gly Phe Leu Met Pro Glu Leu Cys Leu Val Leu Lys Val Arg
 100 105 110

His Ser Glu Asn Thr Leu Phe Ile Ile Ser Asp Ala Ala Asn Leu His
 115 120 125
 Glu Phe His Ala Asp Leu Arg Ser Cys Phe Ala Pro Gln His Met Ala
 130 135 140
 Met Leu Cys Ser Pro Ile Leu Tyr Gly Ser His Thr Ser Leu Gln Glu
 145 150 155 160
 Phe Leu Arg Gln Leu Leu Thr Phe Tyr Lys Val Ala Gly Gly Cys Gln
 165 170 175
 Glu Arg Ser Gln Gly Cys Phe Pro Val Tyr Leu Val Tyr Ser Asp Lys
 180 185 190
 Arg Met Val Gln Thr Ala Ala Gly Asp Tyr Ser Gly Asn Ile Glu Trp
 195 200 205
 Ala Ser Cys Thr Leu Cys Ser Ala Val Arg Arg Ser Cys Cys Ala Pro
 210 215 220
 Ser Glu Ala Val Lys Ser Ala Ala Ile Pro Tyr Trp Leu Leu Leu Thr
 225 230 235 240
 Pro Gln His Leu Asn Val Ile Lys Ala Asp Phe Asn Pro Met Pro Asn
 245 250 255
 Arg Gly Thr His Asn Cys Arg Asn Arg Asn Ser Phe Lys Leu Ser Arg
 260 265 270
 Val Pro Leu Ser Thr Val Leu Leu Asp Pro Thr Arg Ser Cys Thr Gln
 275 280 285
 Pro Arg Gly Ala Phe Ala Asp Gly His Val Leu Glu Leu Leu Val Gly
 290 295 300
 Tyr Arg Phe Val Thr Ala Ile Phe Val Leu Pro His Glu Lys Phe His
 305 310 315 320
 Phe Leu Arg Val Tyr Asn Gln Leu Arg Ala Ser Leu Gln Asp Leu Lys
 325 330 335
 Thr Val Val Ile Ala Lys Thr Pro Gly Thr Gly Gly Ser Pro Gln Gly
 340 345 350
 Ser Phe Ala Asp Gly Gln Pro Ala Glu Arg Arg Ala Ser Asn Asp Gln
 355 360 365
 Arg Pro Gln Glu Val Pro Ala Glu Ala Leu Ala Pro Ala Pro Val Glu
 370 375 380
 Val Pro Ala Pro Ala Pro
 385 390

<210> 7
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 7
 cttgaggatg cggatgtgct

20

<210> 8
 <211> 18
 <212> DNA
 <213> Homo sapiens
 <400> 8
 ccatgggggtg agtgtcct

18

<210> 9
 <211> 18
 <212> DNA
 <213> Homo sapiens
 <400> 9
 aggacactca ccccatgg

18

<210> 10
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 10
 gtatgggaca ggggcagaaa

20

<210> 11
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 11
 tttctaaaga ccattgggag

20

<210> 12
<211> 20
<212> DNA
<213> Homo sapiens

<400> 12

ccatttttaa gtagcggttc 20

<210> 13
<211> 20
<212> DNA
<213> Homo sapiens

<400> 13

aggagagaaa ggtgagccaa 20

<210> 14
<211> 20
<212> DNA
<213> Homo sapiens

<400> 14

gtagatcctg aggttgacca 20

<210> 15
<211> 20
<212> DNA
<213> Homo sapiens

<400> 15

tgtgagcatt tctggccttc 20

<210> 16
<211> 20
<212> DNA
<213> Homo sapiens

<400> 16

tgaagacgcc agagaagcag 20

<210> 17
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 17
 gcctcacaaag tgcagacct

20

<210> 18
 <211> 18
 <212> DNA
 <213> Homo sapiens
 <400> 18
 agaaggggtgg tgaagact

18

<210> 19
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 19
 cttgggttaga gaggatgggc

20

<210> 20
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 20
 gcccatcctc tctaaccaag

20

<210> 21
 <211> 15202
 <212> DNA
 <213> Homo sapiens
 <400> 21
 gatccgagct caattaaccc tcactaaagg gactcgactc gatccttaaa atattcatat 60
 ctctcggaca acctgtggcc atagtgcctg actgtaaacc aaaaggggtt gccctttgcc 120
 gtgtagccca gccctgggtgc tgcctgccct cgcgggtgtct gtgcacctgc cactgatgctg 180

aggccattc gtotgccac tatcttagcg ttttcaaagg gctttcacot ctgaacccag 2100
 gcatcctcgg agatgagtag gtgaagcagg tctcatgagc gtgtctgctg gcccgggccc 2160
 cacggaagag gggagggtgt gccgtcccga gtggagccga ggctcgggac acgcaggaaa 2220
 ggacgcgcgc tgcccgggct cctggagacg cagaactctg tgtgaggtct tgggaaaaa 2280
 gttaacccgc atgttttaag agccagaaaa acattccacc cccttgacct ggtaacccca 2340
 ctgttgggga ttttctctta gagggataag ataccgggaa ggggaggtga aatgctcacc 2400
 actgccaaaa cacgggctgc aactgcaaca tcggaggatg agaggagagag tcggctgttg 2460
 tgcagaatgc tcagcagccc tcccagcagg gacaggaaga ctgggcagga agaggggaga 2520
 agcattcaag ttaaggcaaa aggcccaacg cagagcagca cactgaggtc acacctgtga 2580
 gatgtggaag agaattcctg agcgtggagc gatggggtta ggtgccagga tgattgccca 2640
 ttttgcttct gtcagactct tgactaagga tttctggttg cattttatta cataaaagcc 2700
 agggaggtta tatcacggtg agaaagcttc cctgacgcgc cctcctgtag cgcagccaa 2760
 cgagcctgtg gaggtaccat atgactgtag gcctctgggg acagggagct gcatctgctt 2820
 ctcaaggcca gggacacagc catttctgcc agcatctgtt gatcagttag tgagttagtg 2880
 ggcaggtaga gcaggagcca gtgaagagca ggccttgat ggggtgggat gcaccatgtc 2940
 cccaggctgc agctgcaggc agccccccac attgtcggag aagcctctgc accagctcag 3000
 cccctcctc actccctctg tgccctgggg acactctgca gaggggcact ctgcagctcg 3060
 tcccgcgat cgctggactt ctggacatgg cctccagatt tgcaactctt aaataaatct 3120
 gcagtggatg totttgtgtg cacctctctt tccttttggt gagaacagc aaagatcgga 3180
 cccctaagga ctctctgat gtctccgctc tatccgtga gtgcctttc tgaccacttg 3240
 tttgtacagg ccacggtcca ggacgggagc agatagactg tccctgtccc tgtccacatt 3300
 tcttggtcc aaacagggtt tgtgggaggt agtggcaaaa ggtgttggtc tttttctcac 3360
 tgatttgag gcctcccggt gtgttttttc agccgcgtgt tctgggtct tgcctggatg 3420
 gacagggttt ttaagcgtg gggagcagct ttgctgacca tgccgtgtgc ttccagcctg 3480
 attcccaga agggagcgtg cttgcgaagg aactggcact cgggcctgcc tgaagggggc 3540
 gctgtccaga cacaccagc ctcccgctgt ggcaggcgtg gtcggagcca tggatgattg 3600
 tgaccaatag ggggtgctgc cagagttgat tgtccagcca ggcccagggg ctgagaggag 3660
 gctgtgtgga gaggtgtta ggagccaggg ctcggtcagc tgagttcgca tgccagcttc 3720
 ctgactgtgg gacctcaagc aactgttagc cctctgaag ctgttttctc aactgtgaag 3780
 tggacgcacc ctacttcatt gattctaaga ggcacgcat tccacctgtg gacttctctg 3840
 aaactgaggt gcgtctttca gtcagtggcg tctcatagtc gctgtcagcc agctggtatt 3900

cgagatggag tcgtggaaaa cccgtggaca ccttcgcta ggaccaagat ggagccacct 3960
 gccgcatcct agatttgatg aaatgtggta aataacgaga gccatgcatg agcgaatgct 4020
 ggggaggcgc ttggcactac ccagagctcc acagagggtg tcgatgaggg ctgcccttcc 4080
 ccacatcctt agtaggggtt tcaacatgac ccagactgtg cccctgggga gcttggagcc 4140
 atgcgggagg atgagccatg tgctggagga gaacagggtg ggtgggtgtg gggcttttgt 4200
 agactgtcta gagcagagaa ggtctgcagt ggaggtggtg tctgaggtga atctogaagg 4260
 tgaataggag ttgaacgtta gcaggcagag ggtggattgc aggagagcag cggcctgggc 4320
 aggtgcccat cgtggcccat cagggtgctt catgcatgac tgtgtgcttg ccacccctcc 4380
 tgctgccta cccctgctg cttcgcctca tgggggctt tgagcttggg cccacccgcc 4440
 tgctcgctt gtgggcagag gaccaggtc gtgtgagttg tctgtcccg gggagcagct 4500
 gagcttgtcc gggttcctcg acctgtgggg cttcagagga cttcgggtca tttcaatggg 4560
 ctgtggcgat gctggctgtg gaggtagcct agggctcctg tagccttcag tgagactggc 4620
 ggcccgatgc ccagtgttca ccctgctggc ggcagtcagg aacatgttca caaagcttta 4680
 cttcaagtgg tctagaggtg atctgaggtg gagtaacagg tccagatagg ctacgttcat 4740
 aaaacagctt cagcggggtt taggaacct gtgcatttac gggacgcagt gggtcagagt 4800
 gctgctgtcc gtgggaggtg gccccagggc aggtcagttg gcacgtcctg tggtaagtgg 4860
 gactgtggat gtgggctcag gctggactca gcagccctgc tggataccaa ggcctgcaag 4920
 ggctggcccc ctggtgaatt gtcccgctgc ctgtgtatct atgagtcctg cagagatgac 4980
 aaatcagggg acggggtoat gtctagtca cgtctgggaa atgtctccag gagtgaacac 5040
 atttcaggct cttgatggat gtacctcaa actctctctt ggatgggtgg gccagcttgc 5100
 atgcctgtgc cggcctctgc ccagcgaggt cagggccagg ccacacagtc agtctgactt 5160
 tggcagaagt tgagaggcaa cactgtctc ttgtttcagc ttgcctttct ttgtgtactt 5220
 ctgagagcga gcattctttt catgttctat ccgctggccg ttcttctgag gaatgtctgt 5280
 tcaagtcctt tgcagtctgt taatgaggtt tcaaaccttc cctcattttt gtaatgtgta 5340
 agaacttttt ccagactagc gatataaat cttgtcaaat attgcaacaa cttttctcat 5400
 ttcatctggt tttaactctat cctgggtttt aaaaaatgtg tctgtggaag tttaattttt 5460
 atgtagtca atctcagttt ttttccattg catttattct cagaatgctt ctccctgcc 5520
 tgagattaga taagcagtc tttgttctt cttgagttat ttgagattt cagttttaac 5580
 attttctctt ataatccatg tggctgggtt ttgggatctg gctaaccccc gccatgccag 5640
 tagcctgagg ggccagccc cacttgttga acagccgctc tcccccccc accaccctg 5700
 cctgcctgcc caccgcctt ggtctctcca ggaatcatgt tcgttcagga ggaggccctg 5760

gccagcagcc tctcgtccac tgacagtctg actcccgagc accagcccat tgcccaggga 5820
 tgtttctgatt ccttgaggcg catccctgcg ggacaggtaa tgcctcttcc ccgtctcttg 5880
 ggaccataca tctgtgggtg gactcttctg cttgggggtg tgtgcagtag gaagtggcct 5940
 agctggagct gaggcagatg ctccagggtt ttggcgctct ctgctttgcg ccacggtctt 6000
 tctcttggac ctgtctctgg ttgagtgtct tctgacaaa cacagtgggt aagggtttat 6060
 tttcagcctc cctccttccc tccccaccc accttggttg atgggaacag gcagttctct 6120
 gtcactgggc ccaggggcac aggggggcag gtggagaggg tggcccttga cctgtgagc 6180
 aggcctccct ggggaaggca ttcaaaaga cctcgtgca gggccttgtt tgggtttctt 6240
 ctctgtttcc tggcaccctc ggagccactc ggcgccttcc cgcagtgcac cctggtggtc 6300
 tgggaaacag tctcactctg gcgcctctc tgtggtgtt actgagagtt ctggggcccc 6360
 ttctttgtc ctgaggaag acaggaggaa agcaagggtg cttgctgtgt gcttcgcaaa 6420
 tgtgcttggc gctggggcct cctccagcc ccatctctgc agcagcaaaa ggttatggcc 6480
 ttgtgacact gggacagttt gcagagtctt tgtctgtcct cagtaactca cagtattctg 6540
 ccatcaccct ttccagggtc acacagcaag agattcccaa gccctagtta tccccagtg 6600
 cacagagacc attgggaggg acttgccagg gctgtgtcca ctgctggcca gttagggtcg 6660
 gaccaaatth gttagactgtc tacctggacc cttgcgtggc acaaggagca gttagatgct 6720
 ggatccctgg agagtggcga gaggtctctg ccttaggttg cgagtgggaa tccagccct 6780
 gctgtgtgct ggtgggataa ccaagtgggt ctctgcctt gggctccaga gtgggcccc 6840
 gggctccaga gtgggtccca gggtagagcg tggggatggg gagcctctc agggcggtga 6900
 tggagggcag aatgccaccg tcagggtctg gcaaccagta aatggctggg gctggctgca 6960
 gtagggtggg actgactgtg tttctttctc catcaggcag cttccgatga tttaagggac 7020
 gtgccaggag ctgttggtgg tgcaaggtaa ggaagaggtt ggaaagggac ctgggcctgg 7080
 ccacacagcc ttatgcacac acaactgctgt gggccagggg ttggccagta ggttttttta 7140
 aaaatccgtt cacagaaggc ctatagaact atttcttct ctaagagac acagatgaga 7200
 tggacttttc aatctgtttc caaattctaa taactaaact ctgctcagca catgttgccc 7260
 tacaccaggg gttggcaaat caaggcctgt gtgtggccca cagcctggga gctaagaatg 7320
 acagttacat tctttttctt ttttttgaga ctgagtctcg ctctgtcgcc caggctggag 7380
 tgcagtgccg tgttcttggc tcaactgaac ccccgctccc cagattaatg caattttctt 7440
 gtctcagcct cagccttctg agtagcccg accacaggcg cagccacca cggccaacta 7500
 attttttata ttttttagtg agacagagat tcacatgtg gcctagctgt tctgaactc 7560
 ctgaactcca gtgactccac aacctcggct tcttaaagta ctggaattac aggcattgagc 7620

caccgcgcct ggctagaata acagttactt tttttttctt tgagactgag tcttgctttg 7680
 tcacccaggc tggagtgcag tggcacgac tcagctcgct gcaacctccg cctcccggt 7740
 tcaagcgatt cttctgcctc agccacccaa ggtgcccgcc accacacctg gctaattttt 7800
 ctgtttttag tagggacagg atttcgcat gttggacagt tacattctta aagggtgct 7860
 gaagatcgta tggacatggt agcccataaa tcccaaatg tgtactctga ccttttacag 7920
 aagcttacta actcccactc tacatgtgag ggctgcggtg gccaaagaaga gctggaattt 7980
 aagtgtgaag gtcctaagac ctgccccagc ccacttccct gccccggagg ccacaggggg 8040
 tgacaagtag attcatgccc tggagtgttc cttctctccg gggcttatgg cagcaactga 8100
 atgacttaga agtccatggg agtgctttct gttgtgggaa ctggtgtggt ctgggcatag 8160
 ctgtgccagg cacctatggt ccaagcccct agaagcatag actctgacca aactggcgac 8220
 ccagccttcc agcaggcgag actggctccc accagggccc tcatcctggg aactgacttg 8280
 gccatgtggg aggcttgagg gacccatggg ttggtttctc agggctcagg tgtagcagt 8340
 ggctccagat gtggcagggt ggaggtggga ggggccctcc ccagcatgac actgacctgg 8400
 cctctccctg cacagcccag aacatgccga gccggaggtc cagggtgtgc cggggtctgg 8460
 ccagatcctc ttcctgcctc tcactgcat tggctacacg gccaccaatc aggacttcat 8520
 ccagcgctg agcacactga tcgggcaggc catcgagcgg cagctgcctg cctggatcga 8580
 ggctgccaac cagcgggagg agggccaggg tgaacagggc gaggaggagg atgaggagga 8640
 ggaagaagag gaggacgtgg ctgagaaccg ctactttgaa atggggcccc cagacgtgga 8700
 ggaggaggag ggaggaggcc agggggagga agaggaggag gaaggaggag atgaaggagc 8760
 cgaggaggag cgctgtgctc tggaaatggc cctggggcgg gacgaggact tctgtctgga 8820
 gcacatccgc atcctcaagg tgcgtgtgtg cttcctgac catgtgcagg gcagtatccg 8880
 ccagttcgcc gctcgccttg tgcctaccga cttcggcacc gcagtcttcg agatcccgca 8940
 ccaggagtct cggggcagca gccagcacat cctctcctcc ctgcgctttg tctttgctt 9000
 cccgcatggc gacctcacgc agtttggtt cctcatgccg gagctgtgtc tgggtgctca 9060
 ggtacggcac agtgagaaca cgctcttcat tatctcggac gccgccaaac tgcacaggtt 9120
 ccaocgggac ctgcgctcat gctttgcacc ccagcacatg gccatgctgt gtgacccccat 9180
 cctctacggc agccacacca gcctgcagga gttcctgcgc cagctgctca cctctacaa 9240
 ggtggtctgc ggctgccagg agcgcagcca gggctgcttc ccgctctacc tggctctacg 9300
 tgacaagcgc atggtgcaga cggccgcggg ggactactca ggcaacatcg agtgggcccag 9360
 ctgcacactc tgttcagcgc tgcggcgctc ctgctgcggc cctctgagg ccgtcaagtc 9420
 cgcgcgcatc cctactggc tgttgctcac gccccagcac ctcaacgtca tcaaggccga 9480

ctccaacccc atgccaacc gtggcaccca caactgtcgc aaccgcaaca gcttcaagct 9540
 cagcgcgtgtg cgcgtctcca cegtgtctgt ggaccccaca cgcagctgta cccagcctcg 9600
 gggcgcccttt gctgatggcc acgtgctaga gctgctctgt gggtagccgt ttgtcactgc 9660
 catcttctgt ctgcccacg agaagttcca ctctctcgc gtctacaacc agctgcgggc 9720
 ctgcgtgcag gacctgaaga ctgtggtcat cgccaagacc cccgggacg gaggcagccc 9780
 ccagggtctcc ttgcggtatg gccagcctgc cgagcgcagg gccaggtgag atcaagcaca 9840
 gctctcaggg gccccggggg caggggtctg gcatgtgtgt gatctcagca tctgcggcta 9900
 gtgtgggctg ggagttgctg cgagagctgg gcccccctcc cctgcccct cgccccccc 9960
 gggcctccct ctacatcacc acccagggtt tggtgccagg ctgctcctta tctcagtgt 10020
 gtagaagaag cccaggaaag ctgtctcttc acaaaatggg ttggcccagc ctcttgccc 10080
 ccatgaaggg caggccaagg gggctgcccc accttgctct gccagtgagg agagcaaacg 10140
 gctgcagcac accgaggcca ggagagctgt caccctggct gctgtgctcc tctgggcccc 10200
 agcatggcct ctggggcacta cctcctccag ggtcacagtc ccacggatgg ctctgtgggc 10260
 caggatctgc cttaggcttc acccacctca acatcttctg gtgtgtgtca ggctggtctc 10320
 aaactttggg ctcaaaacat cctccgcctc agcctcccaa agtgctggga ttacagacat 10380
 gagccaccgt gcccgccctg gctgttctgt tctccaatag agaagctggt ggaagtcctc 10440
 agtaaccag aggtgatgtg tgatgcacac agtctcctca ctctgaagct gcacatgcga 10500
 tgtgaatctt catttggggg ccgctgttaa tatggtgttt ttgggggat acagcaatga 10560
 ccagcgtccc caggaggctc cagcagaggg cctggccccg gccccagtg aagtcaccgc 10620
 tccagccctc gcagcagcct cagcctcagg cccagcgaag actccggccc cagcagaggg 10680
 ctcaaactta gctttggtcc cagagagagc gccagtgga gctccagccc cacccccagc 10740
 cgaggccctc gcccagtagc cgagtgcagc cctcatccag gccacctcg aggagaatca 10800
 gatccccctg caactgctcg cctgccctgc gctccggcac gtgcggagcc tgcggggcag 10860
 cgccatcctc gagctcttcc acagcagcat tgctgagga gcggcccggg tgtgggtgcc 10920
 agctatggca cggccagctc tgaggggcag gccaaagctg gcttcaggtc agcctcaggt 10980
 ccttggaact cctgatgtgc ggagtctca gctgagctgc tcacagcttt gaggacctgg 11040
 gcagtagggt cctgagttgc cctccctggc catttctgt gtgtcaccac ctctctgtgc 11100
 acttccagcc ccaggtagac tccccaccaa cagccatctc ccacccctct ctctctctct 11160
 gccttgaagc atacggatct attggtgagc caagaggggc ttcccatgtc tctctgtgga 11220
 agctgtgggc atgtccctgg tatgtgcagg ttgctagggt ggtggagctg acaggaggcc 11280
 ccccgctctc aggttgaaaa cgaggagctg aggcacctca tgtggtctc ggtgggtgttc 11340

taccagaccc cagggtcga ggtgactgcc tgcgtgctgc tctccacca ggtgtgtac 11400
ttttgtctcc acgacggcct ccgccgtac ttctcagagc cactcgaggg taggcacagg 11460
gcctgtctgg gctcaggagc ttggagtgtg tggttggggc aggcctgggg ggtcattctc 11520
tggagccagc tgtgtgctt caggcagcag tcacgcagctt ggtcgcagt ggtcgagagt 11580
tccttgtctg aggaaggag ctgtcatgag ggaggggtcc atggccagat gtgaaccag 11640
aatgcactga gccaggcct ggtgactgct tgggaacagc ctgtgatgag aagggttag 11700
gcagcctttg cccctggggc tgcaacagaa gccctagcca gcgacctggt gactcccctg 11760
agctggaaga ggctcagact ccagagggca ttgcctatgg ggctttgcac ggggtgaagc 11820
caggccagcc aagaggacct gtctctgtcg gatgtgtgc acacctagga accttgtgtc 11880
tgcctgccac cgcctccctc tgcctcttc tccatcacac agatttctgg catcagaaaa 11940
acaccgacta caacaacagc cctttccaca tctccagtg ctctgtgcta aagcttagtg 12000
acctgcagtc agtcaatgtg gggcttttcg accagcattt ccggctgacg cgtgggtgac 12060
cctctgtgtc ttgtctatt tccgggtgaag gccagcatca coagtgggct tccaccttc 12120
gtacgtgggt ggggttatcat agacagttat ctctgtgtgc aagagccact tottaaccgg 12180
ggtgggagga agcagcttca ggaactgctg agagagcaga actcacgctc cagggctcag 12240
agcaggaggt aggggtgtcg gcaagcgctg gcccgagacg aagcagagtg ggccctggtc 12300
tcgggcagga tgtttctgac tcacatttcc tgaggagaga aagctaaagct ctttgccata 12360
tgtctctgtc tccccttcca gaaaaatgcc tcagctcttc cggcctgaag gaatggcctc 12420
ctcccgggcc ccattgattt tctctgtgtg ggcctctctg gccctggcct ctgggtgtag 12480
gcttgctagg gactcgggtt ggctctaaag gcaggggata gggctgggga gcgcggcct 12540
gtggccctga ccagccctct ctctgtcagg ttccaccccg atgcaggtag tcacgtgctt 12600
gacgggggac agctacctga cgcactgctt cctccagcac ctcatggtcg tctgtctctc 12660
tctggaacgc acgcccctgc cggagcctgt tgacaaggac ttctactccg agtttgggaa 12720
caagaccaca ggtacccctg tctagctcag gctgcagaca ggtgcctgg acagacgtca 12780
tgggcccagc ggtggtctc tgtgccccag aacctctct gctctatgt ctctctttc 12840
tcacttagct ggccagggtt ttatgtgggg cttttcgatg gcagagtctc cactccagca 12900
gtccctcaac catctggcag acacatctcc agtgctgtct ttgggtctct ggctgtggg 12960
ccccacactt ggagcactct ctctctgctg tctcatgccg gggctctctg gttggcttg 13020
ggcccttggt gctccagcc ccaccagggg ccggttccag gctatagccc aggtggcctc 13080
tctctgcagg gaagatggag aactacgagc tgatccactc tagtcgctc aagtttacct 13140
acccagtgta ggaggagatt ggggacctga cgttactgt ggcccaaaag atggctgagc 13200

cagagaaggc cccagccctc agcatcctgc tgtacgtgca ggcottccag gtgggcatgc 13260
caccctctgg gtgctgcagg ggccccctgc gcccgaagc actcctgctc accagctccg 13320
agatcttctt cctggatgag gactgtgtcc actaccact gcccagagtt gccaaagagc 13380
cgccgcagag agacaggtag cggctggacg atggcccgcg cgtccgggac ctggaccgag 13440
tgctcatggg ctaccagacc taccgcagg cctccacct cgtcttcgat gacgtgcaag 13500
gtcatgacct catgggcagt gtcacctgg accactttgg ggaggtgcc a ggtggcccg 13560
ctagagccag ccagggccgt gaagtcagtg ggcaggtgtt tgtccccagt gctgagagca 13620
gagagaagct catctcgtg ttggctcgcc agtgggagcg cctgtgtggt cgtgagctgc 13680
ctgtcgagct caccggctag ccagggccac agccagcctg tctgttcagg cctgaocct 13740
actggggcag ggcagcaggc ttttgtgttc tctaaaaatg ttttatctc cctttggtac 13800
cttaatttga ctgtctcgc agagaatgtg aacatgtgtg tgtgttgtgt taattctttc 13860
tcatgttggg agtgagaatg ccgggccctc cagggtctgc ggtgtgtgt cagcctccca 13920
caggtgttac agcgtgcac accagtgtcg tgtctgtgt tgtgggaccg ttgttaacac 13980
gtgacactgt gggctcgtac ttctcttcta cagctcctt cctgaagtgt cgagtccagt 14040
cctttgttgc tgttgtgtt gctgtgtgt ttgtgttgg catcttctg ctaaatctga 14100
ggctggtagc agaatgcaca ttggaagctc ccaccccata ttgtcttca aagtggagg 14160
ctccctgat ccagacaagt ggagagccc gtggggcgag gggacctgga gctgccagca 14220
ccaagcgtga ttctgtctgc ctgtattctc tattccaata aagcagagtt tgacaccgtc 14280
tgcatottct aaaccaaggg tcaactggat cgagtgcag gccctatagt gagtctgatt 14340
agagctcgcg gccgcgagct ctatagtcac gctcgacgg ccgccagtgt gatggatatc 14400
tgcagaattc cagcactgc gccggcgtta ctagtggatc cgagctccac agaggtggtc 14460
gatgagggct gccctttccc acatccttag tagggggttc aagatgaccc agactgtgcc 14520
cctggggagc ttggagccat gccggaggat gagccatgtg ctggaggaga acagggtagg 14580
atggtgtggg gctttttag actgtctaga agcaaagaag gtctgcagtg gaggtgtgt 14640
ctgaggtgaa tctcgaagt gaataggagt tgaacgttag caggcagagg gtggattgca 14700
ggagagcagc ggctgggca ggtgcccgag gtggcccatc aggggtgttc atgcatggct 14760
gtgtgcttgc catcttctt gctgcctac cccctgtgc ttctgttcat gggggcgctt 14820
gagcttgggc ccactgctt gctcgtctt tgggcagag acccaagctg tgtgagttgt 14880
cctgtcccg ggagcagctg aactggtccg ggtgtctgaa ctgtggggct caaaggact 14940
ccggggctat ttactgggg ctgtgccgat tctgggggc tgttngaan gtaaaggcct 15000
aaaggggctc cctggttang gccctcaant ttaanaacct ggggccgggg ccggaattg 15060

cccccaantt tgtttcaacn ccccttggcc ttnggcnggg gcaaatttcc anggggaacc 15120
aatggnnttt ccccaaaaaa ggggcntttt taaccntttt ccaaantttg ggnccataaa 15180
aaggggtggan ttcctgaang gg 15202

<210> 22
<211> 1070
<212> PRT
<213> Homo sapiens

<400> 22

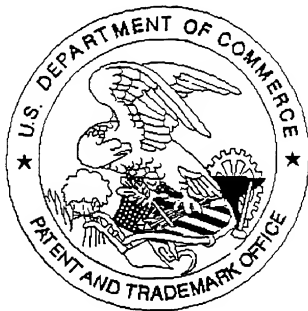
Val	Cys	Leu	Asp	Asp	Thr	Val	Thr	Thr	Glu	Lys	Glu	Leu	Asp	Thr	Val
1				5					10					15	
Glu	Val	Leu	Lys	Ala	Ile	Gln	Lys	Ala	Lys	Glu	Val	Lys	Ser	Lys	Leu
		20					25						30		
Ser	Asn	Pro	Glu	Lys	Lys	Gly	Gly	Glu	Asp	Ser	Arg	Leu	Ser	Ala	Ala
		35				40						45			
Pro	Cys	Ile	Arg	Pro	Ser	Ser	Ser	Pro	Pro	Thr	Val	Ala	Pro	Ala	Ser
	50					55					60				
Ala	Ser	Leu	Pro	Gln	Pro	Ile	Leu	Ser	Asn	Gln	Gly	Ile	Met	Phe	Val
65				70					75					80	
Gln	Glu	Glu	Ala	Leu	Ala	Ser	Ser	Leu	Ser	Ser	Thr	Asp	Ser	Leu	Thr
			85						90					95	
Pro	Glu	His	Gln	Pro	Ile	Ala	Gln	Gly	Cys	Ser	Asp	Ser	Leu	Glu	Ser
		100					105						110		
Ile	Pro	Ala	Gly	Gln	Ala	Ala	Ser	Asp	Asp	Leu	Arg	Asp	Val	Pro	Gly
		115					120					125			
Ala	Val	Gly	Gly	Ala	Ser	Pro	Glu	His	Ala	Glu	Pro	Glu	Val	Gln	Val
	130					135					140				
Val	Pro	Gly	Ser	Gly	Gln	Ile	Ile	Phe	Leu	Pro	Phe	Thr	Cys	Ile	Gly
145				150					155					160	
Tyr	Thr	Ala	Thr	Asn	Gln	Asp	Phe	Ile	Gln	Arg	Leu	Ser	Thr	Leu	Ile
			165						170					175	
Arg	Gln	Ala	Ile	Glu	Arg	Gln	Leu	Pro	Ala	Trp	Ile	Glu	Ala	Ala	Asn
		180					185						190		
Gln	Arg	Glu	Glu	Gly	Gln	Gly	Glu	Gln	Gly	Glu	Glu	Glu	Asp	Glu	Glu
		195				200						205			
Glu	Glu	Glu	Glu	Glu	Asp	Val	Ala	Glu	Asn	Arg	Tyr	Phe	Glu	Met	Gly
	210					215					220				
Pro	Pro	Asp	Val	Glu	Glu	Glu	Glu	Gly	Gly	Gly	Gln	Gly	Glu	Glu	Glu
225					230					235					240

Glu Glu Glu Glu Glu Asp Glu Glu Ala Glu Glu Glu Arg Leu Ala Leu
 245 250 255
 Glu Trp Ala Leu Glu Gly Ala Asp Glu Asp Phe Leu Leu Glu His Ile Arg
 260 265 270
 Ile Leu Lys Val Leu Trp Cys Phe Leu Ile His Val Gln Gly Ser Ile
 275 280 285
 Arg Gln Phe Ala Ala Cys Leu Val Leu Thr Asp Phe Gly Ile Ala Val
 290 295 300
 Phe Glu Ile Pro His Gln Glu Ser Arg Gly Ser Ser Gln His Ile Leu
 305 310 315 320
 Ser Ser Leu Arg Phe Val Phe Cys Phe Pro His Gly Asp Leu Thr Glu
 325 330 335
 Phe Gly Phe Leu Met Pro Glu Leu Cys Leu Val Leu Lys Val Arg His
 340 345 350
 Ser Glu Asn Thr Leu Phe Ile Ile Ser Asp Ala Ala Asn Leu His Glu
 355 360 365
 Phe His Ala Asp Leu Arg Ser Cys Phe Ala Pro Gln His Met Ala Met
 370 375 380
 Leu Cys Ser Pro Ile Leu Tyr Gly Ser His Thr Ser Leu Gln Glu Phe
 385 390 395 400
 Leu Arg Gln Leu Leu Thr Phe Tyr Lys Val Ala Gly Gly Cys Gln Glu
 405 410 415
 Arg Ser Gln Gly Cys Phe Pro Val Tyr Leu Val Tyr Ser Asp Lys Arg
 420 425 430
 Met Val Gln Thr Ala Ala Gly Asp Tyr Ser Gly Asn Ile Glu Trp Ala
 435 440 445
 Ser Cys Thr Leu Cys Ser Ala Val Arg Arg Ser Cys Cys Ala Pro Ser
 450 455 460
 Glu Ala Val Lys Ser Ala Ala Ile Pro Tyr Trp Leu Leu Leu Thr Pro
 465 470 475 480
 Gln His Leu Asn Val Ile Lys Ala Asp Phe Asn Pro Met Pro Asn Arg
 485 490 495
 Gly Thr His Asn Cys Arg Asn Arg Asn Ser Phe Lys Leu Ser Arg Val
 500 505 510
 Pro Leu Ser Thr Val Leu Leu Asp Pro Thr Arg Ser Cys Thr Gln Pro
 515 520 525
 Arg Gly Ala Phe Ala Asp Gly His Val Leu Glu Leu Leu Val Gly Tyr
 530 535 540
 Arg Phe Val Thr Ala Ile Phe Val Leu Pro His Glu Lys Phe His Phe
 545 550 555 560
 Leu Arg Val Tyr Asn Gln Leu Arg Ala Ser Leu Gln Asp Leu Lys Thr
 565 570 575

Val Val Ile Ala Lys Thr Pro Gly Thr Gly Gly Ser Pro Gln Gly Ser
 580 585 590
 Phe Ala Asp Gly Gln Pro Ala Glu Arg Arg Ala Ser Asn Asp Gln Arg
 595 600 605
 Pro Gln Glu Val Pro Ala Glu Ala Leu Ala Pro Ala Pro Val Glu Val
 610 615 620
 Pro Ala Pro Ala Pro Ala Ala Ala Ser Ala Ser Gly Pro Ala Lys Thr
 625 630 635 640
 Pro Ala Pro Ala Glu Ala Ser Thr Ser Ala Leu Val Pro Glu Glu Thr
 645 650 655
 Pro Val Glu Ala Pro Ala Pro Pro Pro Ala Glu Ala Pro Ala Gln Tyr
 660 665 670
 Pro Ser Glu His Leu Ile Gln Ala Thr Ser Glu Glu Asn Gln Ile Pro
 675 680 685
 Ser His Leu Pro Ala Cys Pro Ser Leu Arg His Val Ala Ser Leu Arg
 690 695 700
 Gly Ser Ala Ile Ile Glu Leu Phe His Ser Ser Ile Ala Glu Val Glu
 705 710 715 720
 Asn Glu Glu Leu Arg His Leu Met Trp Ser Ser Val Val Phe Tyr Gln
 725 730 735
 Thr Pro Gly Leu Glu Val Thr Ala Cys Val Leu Leu Ser Thr Lys Ala
 740 745 750
 Val Tyr Phe Val Leu His Asp Gly Leu Arg Arg Tyr Phe Ser Glu Pro
 755 760 765
 Leu Gln Asp Phe Trp His Gln Lys Asn Thr Asp Tyr Asn Asn Ser Pro
 770 775 780
 Phe His Ile Ser Gln Cys Phe Val Leu Lys Leu Ser Asp Leu Gln Ser
 785 790 795 800
 Val Asn Val Gly Leu Phe Asp Gln His Phe Arg Leu Thr Gly Ser Thr
 805 810 815
 Pro Met Gln Val Val Thr Cys Leu Thr Arg Asp Ser Tyr Leu Thr His
 820 825 830
 Cys Phe Leu Gln His Leu Met Val Val Leu Ser Ser Leu Glu Arg Thr
 835 840 845
 Pro Ser Pro Glu Pro Val Asp Lys Asp Phe Tyr Ser Glu Phe Gly Asn
 850 855 860
 Lys Thr Thr Gly Lys Met Glu Asn Tyr Glu Leu Ile His Ser Ser Arg
 865 870 875 880
 Val Lys Phe Thr Tyr Pro Ser Glu Glu Glu Ile Gly Asp Leu Thr Phe
 885 890 895
 Thr Val Ala Gln Lys Met Ala Glu Pro Glu Lys Ala Pro Ala Leu Ser
 900 905 910

Ile Leu Leu Tyr Val Gln Ala Phe Gln Val Gly Met Pro Pro Pro Gly
 915 920 925
 Cys Cys Arg Gly Pro Leu Arg Pro Lys Thr Leu Leu Thr Ser Ser
 930 935 940
 Glu Ile Phe Leu Leu Asp Glu Asp Cys Val His Tyr Pro Leu Pro Glu
 945 950 955 960
 Phe Ala Lys Glu Pro Gln Arg Asp Arg Tyr Arg Leu Asp Asp Gly
 965 970 975
 Arg Arg Val Arg Asp Leu Asp Arg Val Leu Met Gly Tyr Gln Thr Tyr
 980 985 990
 Pro Gln Ala Leu Thr Leu Val Phe Asp Asp Val Gln Gly His Asp Leu
 995 1000 1005
 Met Gly Ser Val Thr Leu Asp His Phe Gly Glu Val Pro Gly Gly Pro
 1010 1015 1020
 Ala Arg Ala Ser Gln Gly Arg Glu Val Gln Trp Gln Val Phe Val Pro
 1025 1030 1035 1040
 Ser Ala Glu Ser Arg Glu Lys Leu Ile Ser Leu Leu Ala Arg Gln Trp
 1045 1050 1055
 Glu Ala Leu Cys Gly Arg Glu Leu Pro Val Glu Leu Thr Gly
 1060 1065 1070

United States Patent & Trademark Office
Office of Initial Patent Examination -- Scanning Division



Application deficiencies were found during scanning:

☐ Page(s) _____ of _____ were not present
for scanning. (Document title)

☐ Page(s) _____ of _____ were not present
for scanning. (Document title)

There are 5 sheets of drawing and 2 substitute drawings

☐ Scanned copy is best available.